

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 23, 2002, 18:42:32 ; Search time 1581.67 Seconds

(Without alignments)  
10492.810 Million cell updates/sec

Title: US-09-453-387A-1  
Perfect score: 1006  
Sequence: 1 taacacgcgtatctcttc.....aatgcttaattataaaaa 1006

Scoring table: IDENTITY\_NUC  
Gapop 10.0, Gapext 1.0

Searched: 1472140 seqs, 8248589755 residues  
Total number of hits satisfying chosen parameters: 2944280

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : GenEmbl:\*

- 1: gb\_ba:\*
- 2: gb\_hcg:\*
- 3: gb\_in:\*
- 4: gb\_om:\*
- 5: gb\_ov:\*
- 6: gb\_pat:\*
- 7: gb\_ph:\*
- 8: gb\_pl:\*
- 9: gb\_pr:\*
- 10: gb\_ro:\*
- 11: gb\_sts:\*
- 12: gb\_sy:\*
- 13: gb\_un:\*
- 14: gb\_vl:\*
- 15: em\_ba:\*
- 16: em\_fun:\*
- 17: em\_hum:\*
- 18: em\_in:\*
- 19: em\_om:\*
- 20: em\_or:\*
- 21: em\_ov:\*
- 22: em\_pat:\*
- 23: em\_ph:\*
- 24: em\_pl:\*
- 25: em\_ro:\*
- 26: em\_sts:\*
- 27: em\_sy:\*
- 28: em\_un:\*
- 29: em\_vl:\*
- 30: em\_htgo\_hum:\*
- 31: em\_htgo\_hum:\*
- 32: em\_htgo\_rod:\*
- 33: em\_htgo\_hum:\*
- 34: em\_htgo\_inv:\*
- 35: em\_htgo\_rod:\*
- 36: em\_htgo\_other:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	1006	100.0	1006	8	COTMYBA	L04497 Cotton DNA-
2	282.4	28.1	1145	8	AF336286	AF336286 Gossypium
3	263.8	26.2	764	8	AT02636	U2636 Arabidopsis
4	254.2	25.3	1152	8	LETMH27	X95296 L. esculentu
5	253.6	25.2	639	8	AF207991	AF207991 Arabidops
6	248	24.7	1003	8	AF062860	AF062860 Arabidops
7	242.4	24.1	897	8	AF062859	AF062859 Arabidops
8	234.6	23.3	844	8	AT026937	U26937 Arabidops
9	217	21.6	1151	8	AF034134	AF034134 Gossypium
10	216.4	21.5	1284	8	HVMXB1	X70877 H. vulgare m
11	206.8	20.6	1176	8	D88617	D88617 Oryza sativ
12	205.6	20.4	1102	8	HVMXB2	X70876 H. vulgare m
13	201	20.0	1066	8	D88618	D88618 Oryza sativ
14	194.2	19.3	1335	8	ATCM4GENE	X90382 Arabidops
15	189.2	18.8	1118	8	AMMIXTA	X79108 A. majus mix
16	189	18.7	1337	8	LETMH16	X99210 L. esculentu
17	187.8	18.7	1493	8	AF062916	AF062916 Arabidops
18	187.4	18.6	1135	8	AF401220	AF401220 Fragaria
19	186.6	18.5	984	8	AF175990	AF175990 Arabidops
20	186	18.5	1364	8	AB005889	AB005889 Arabidops
21	185.6	18.4	795	8	AB029160	AB029160 Glycine m
22	183	18.2	966	8	AF249310	AF249310 Arabidops
23	183	18.2	991	8	AF336285	AF336285 Gossypium
24	182.8	18.2	795	8	AB029161	AB029161 Glycine m
25	181.8	18.1	1588	8	PHMXPB31	Z13996 P. hybrida m
26	180.2	17.9	1148	8	AF386932	AF386932 Arabidops
27	179.2	17.8	960	8	AF175991	AF175991 Arabidops
28	178.6	17.8	1002	8	AF371983	AF371983 Arabidops
29	178.6	17.8	1038	8	AF249309	AF249309 Arabidops
30	176.2	17.5	1220	8	AF161711	AF161711 Pimpinell
31	176	17.5	1332	8	AWA6292	AJ0006292 Anthrithn
32	175	17.4	1252	8	ATMIXTA	X99809 A. Thaliana
33	174.8	17.4	1032	8	LETHM18CE	X98308 L. esculentu
34	174.4	17.3	819	8	AB029165	AB029165 Glycine m
35	174.2	17.3	1022	8	AF062869	AF062869 Arabidops
36	174.2	17.3	1462	8	PM039448	U39448 Picea maria
37	174.2	17.3	1612	8	AY035145	AY035145 Arabidops
38	172.4	17.1	1101	8	AF175998	AF175998 Arabidops
39	171.6	17.1	1292	8	AF062872	AF062872 Arabidops
40	171	17.0	981	8	AF370613	AF370613 Arabidops
41	170.6	17.0	113800	8	AC006567	AC006567 Arabidops
42	170.6	17.0	197394	8	ATCHRIV27	AL161515 Arabidops
43	169.8	16.9	1384	8	AF062887	AF062887 Arabidops
44	169.2	16.8	1017	8	AF175992	AF175992 Arabidops
45	168	16.7	828	8	AB029162	AB029162 Glycine m

## ALIGNMENTS

RESULT 1  
LOCUS COTMYBA 1006 bp mRNA  
DEFINITION Cotton DNA-binding domain mRNA.  
ACCESSION L04497  
VERSION L04497.1 GI:437326  
KEYWORDS  
SOURCE Gossypium hirsutum (cultivar Acala Sf-2) 3-day pre-anthesis ovule  
CDNA to mRNA.

## ORGANISM

Gossypium hirsutum  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
Rosidae; eustosids II; Malvales; Malvaceae; Gossypium.

REFERENCE 1 (bases 1 to 1006):  
WILKINS, T.A. and LU, C.-C.

## JOURNAL

Unpublished

## FEATURES

location/qualifiers  
1..1006  
/organism="Gossypium hirsutum"  
/cultivar="Acala Sf-2"

Query Match	Score	DB	Length
Best Local Similarity	80.3%	Pred	No. 2.4e-63;

	Matches	331;	Conservative	0;	Mismatches	81;	Indels	0;	Gaps	0;
Oy	39	ttgattgaattt	gtaactgattggaagacatacactt	gttgttgyaaaggtctataccaaca	98					
Db	49	TTTTCTTTGT	TACTGATTAAT	GGAAAGGTCCCTTGTGTGAGAAAGCTATACGACA	108					
Oy	99	aaggtgctctga	ccaagaagagatacaacgcctatcaatactacatc	cgtygtccatgtyt	158					
Db	109	AAGGTCGTGG	CTAAAGAAAGATGAT	TCCGCTCATAGCTTACATCCAGAGCCCATGCTG	168					
Oy	159	aagctgctg	gcgtccctccccaagctgctgcggcgtctgatagtgtgtaagatgtga	218						
Db	169	AAGGTGCTGG	GTGTACCTCCCTAAAGCTGCTGGGCTTCTCCGCTGTGGCAAAAGTTGTA	228						
Oy	219	gattaaatgataa	actacttgaagcctgatacttaagagagaaattcactgaaag	278						
Db	229	.GACTTCGTTGG	ATCACTTAACACCTGATCTAAAGTGCGCAATTCACGTGAAGAAG	288						
Oy	279	aagatgaactata	ctacaagctcaagcttactcttggaacaaatggtattgattgcg	338						
Db	289	AAGATGACCT	CTAATTTCAAGCTGACACGCTTCTGTGTAAAGGTCTCTTAATAGCGG	348						
Oy	339	gaagattccagga	agaaacagataatgataaagaactactggaacacacacatcaaa	398						
Db	349	GGAAGTTACCG	AGGAACACGATTAATGAGATTAAAGATTAATCTAGGAACACGCAATPADAAGAA	408						
Oy	399	gaagactataag	cagaggaatttgataccaacaactcatcgtctctcaatca	450						
Db	409	GGAAGCTTATG	AGCAGAGAGTATTGATCCAGCACTACACAGGCCACTCAATGA	460						

RESULT	3
ATU026936	
LOCUS	
DEFINITION	Arabidopsis thaliana clone myb6 DNA-binding protein mRNA, complete cds.
ACCESSION	U26936
VERSION	U26936.1
KEYWORDS	GI:1020154
SOURCE	thale cress.
ORGANISM	Arabidopsis thaliana Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsids. 1 (bases 1 to 764) Ll,S.F. and Parish,R.W. Isolation of two novel myb-like genes from Arabidopsis and studies on the DNA-binding properties of their products Plant J. 8 (6), 963-972 (1995)
JOURNAL	Medline
MEDLINE	96132152
REFERENCE	2 (bases 1 to 764) Heazlewood,J.L. Direct Submission Submitted (12-MAY-1995) Joshua L. Heazlewood, Botany, La Trobe University, Melbourne, Victoria 3083, Australia
AUTHORS	Location/Qualifiers
TITLE	1..764
JOURNAL	
FEATURES	
SOURCE	

```

CDS
    /db_xref="taxon:3702"
    /clone="myb6"
    15..725
    /codon_start=1
    /product="DNA-binding protein"
    /protein_id="AA98761.1"
    /db_xref="gi:102015"
    /translation="MGRSPCCCEKAHTNKGAWTKEDQRLVDYIRNHGECWRSLPKSA
    GLRSGCSRLRIWNLPLDRCGNFTDDEOIIILKLSLGNKMSLTLAGRLPRTDNL
    EIKYNNWTKIRKLLSHGIDPQTHRIESTSVSYVPIQNDAAVEKSPFNLAVKPK
    TENSQDNGASTSGTTDEDLQNGECYISDNSGHILKLNLDLTLLFGSMNSGRIVGVSS
    ADSFPCDPIVEARLSLL"

```

	Query Match	26.24;	Score 263.8;	DB 8;	Length 764;	
	Best Local Similarity	79.24;	Pred. 1.7e-58;			
	Matches 313;	Conservative 0;	Mismatches 82;	Indels 0;	Gaps 0;	
QY	59	atggagcatcaacctgtgtgtgtaaaaggctctataccaaanaaagtgcgtgacccaagaag				118
Db	15	ATGGGAATATCTCCCTGTGTGTGTAATAAAGCTCACAAACAAAGAGCTTGACTAAGNA				74
QY	119	gaagtaagagcctcatcaactacacatccgtgccatggtgaagagtgctgcgagctccccc				178
Db	75	GAAGATCAACGCTCTGTGATTAATACCGTAATCACGGGAAGGTTGTGGGCTTCTT				134
QY	179	cccaagctgtcgtgctgcttagatgtgtgaagatlcagatltagaatgataactac				238
Db	135	CCTAAATCCGGGAGATTGTGTGGCTGTGGTAATAAGTTAGATTGAGATTAATTATAC				194
QY	239	ttgagagcctgatctctaagaagagaattcaactgtaagaagaagatgagctctcaatcaag				298
Db	195	CTTCTCCTCGATCTTAACGCTGATTAATTTACTGTGATGAAGATCAATCTCATCAAA				254
QY	299	cttcacagttactctggaacaacatgycatctgattcgtggaagatltaccagaagaaca				358
Db	255	CTCCATTAAGTAACTGGTAAACAAAGGTCATTTGATGACGTGAAGATTACAGGAAGACA				314
QY	359	gataatgataaagaactctggaacacacacatcaaaaagaagcttataagcagaaga				418
Db	315	GATAACGAATTAAGAAATTAATGAGACACTCATTAATTAAGAGGAACCTTTAGTACGGT				374
QY	419	attgatccaaaactcatcgtctctcacaac	453			
Db	375	ATTGATCCCAACATCATCTGTCAGTTTAACGATC	409			

RESULT	4
LOCUS	LETHM27
DEFINITION	L.esculentum mRNA for THM27 protein.
ACCESSION	X95296
VERSION	X95296.1 GI:1167483
KEYWORDS	myb-related gene; THM27 gene; transcription factor.
SOURCE	tomato.
ORGANISM	Lycopersicon esculentum
	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Asteridae; eunastids I; Solanales; Solanaceae; Solanum;
	Lycopersicon.
	1 (bases 1 to 1152)
REFERENCE	Lin, Q., Hamilton, W.D. and Merryweather, A.
AUTHORS	Cloning and initial characterization of 14 myb-related cDNAs from tomato (Lycopersicon esculentum cv. Alisa Craig)
TITLE	Plant Mol. Biol. 30 (5), 1009-1020 (1996)
JOURNAL	96270378
MEDLINE	2 (bases 1 to 1152)
REFERENCE	Lin, Q.
AUTHORS	Direct Submission
TITLE	Submitted (23-JAN-1996) Q. Lin, John Innes Centre, Colney Lane, Norwich, NR4 7UH, UK
JOURNAL	location/Qualifiers
FEATURES	1..1152
source	/organism="Lycopersicon esculentum"
	/cultivar="Alisa Craig"
	/db_xref="taxon:4081"
	/dev_stage="3 weeks old seedling"
	/tissue_type="hypocotyl"
	111..932
gene	/gene="THM27"
	111..932
	/gene="THM27"
	/note="myb-related"
	/codon_start=1
CDS	/product="transcription factor"





TITLE Direct Submission  
JOURNAL Submitted (04-May-1998) Abt. Biochemie, MPI fur Zuchtungsforchung,  
Carl-von-Linne-Weg 10, Koeln D-50829, Germany

FEATURES  
SOURCE Location/Qualifiers  
1.1003  
/organism="Arabidopsis thaliana"  
/cultivar="Columbia"  
/db\_xref="taxon:3702"  
/chromosome="IV"  
/map="near marker ve025"  
1.1003  
/gene="MYB4"  
/allele="COL"  
39.887  
/gene="MYB4"  
/note="MYB4: R2R3-MYB factor family member"  
/codon\_start=1  
/product="putative transcription factor"  
/protein\_id="AAC83582.1"  
/db\_xref="GI:3941412"

gene  
1.1003  
/gene="MYB4"  
/allele="COL"  
39.887  
/gene="MYB4"  
/note="MYB4: R2R3-MYB factor family member"  
/codon\_start=1  
/product="putative transcription factor"  
/protein\_id="AAC83582.1"  
/db\_xref="GI:3941412"

CDS  
/translation="MGRSPCCERKATNKAWTEEDERLVAITKANGEGMRSLPKAA  
GLTRGKSCRLKWLINLPDLKGNFTEDDELITKLSLGNKWSLIAGLPERITN  
EIKWYNTHIRKRLINRGIDPTSHRPIQSSASQDSKPIOLEVTSMTINISFSAK  
VETFHESISFPGSEKISMLTFKEKDECPVOEKFDPDLNLEISLPDVRLOGHCK  
STPRCFKCSLGMINGMECRGMRCDVVGSSKSGSDMSGDFELGAKKETTSILGF  
RSLKWK"

BASE COUNT 326 a 200 c 234 g 243 t  
ORIGIN

Query Match 24.7% Score 248; DB 8; Length 1003;  
Best Local Similarity 76.2% Pred. No. 2.5e-54;  
Matches 305; Conservative 0; Mismatches 95; Indels 0; Gaps 0;

QY 54 aactgtaggagcagatcactctgtgtgaaagagctcattacaacaagagtcctggacca 113  
DB 34 AAGATGAGGAGGTCACCGTGTGTGAGAAAGCTCAACAACAAAGAGCATGAGACA 93  
QY 114 aagaggaagatcaagcgtcatcaactacatccgtgtcattgtgaaagctgtggcgtt 173  
DB 94 AAGAGAGAGACAGAGAGCGCTGCGCTACATTAAGCTGAGAGAGCGCTGAGAGAT 153  
QY 174 cccctcccaagcgt 233  
DB 154 CTCTCCCAAGCGCGGAGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 213  
QY 234 actactgagcgt 293  
DB 214 ACTATCTCGGCGCTGACCTTAAGCGTGAAGTCAACGAGAGAGAGAGAGAGATCA 273  
QY 294 tcaagcttcacagttacttgaagaacaatgtgtgtgtgtgtgtgtgtgtgtgtgtgt 353  
DB 274 TCAAGCTTCATGAGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 333  
QY 354 gaaacagataatgagataaagaactactggaacacacacacacacacacacacacacac 413  
DB 334 GAAACAGATTAACGAGATTAAGAGACTATTGGAACACGCAATATGAGAGAGAGAG 393  
QY 414 gaggaattgattccacaactatcgtctctcattcaataaac 453  
DB 394 GAGGAGTTGATCCAGAGATTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGATC 433

RESULT 7  
LOCUS AF062859 897 bp mRNA PLN 09-DEC-1998  
DEFINITION Arabidopsis thaliana putative transcription factor (MYB3) mRNA,  
complete cds.  
ACCESSION AF062859  
VERSION AF062859.1 GI:3941409  
KEYWORDS  
SOURCE thale cress.  
ORGANISM Arabidopsis thaliana

REFERENCE  
AUTHORS  
1 (bases 1 to 897)  
Kranz, H.D., Denekamp, M., Greco, R., Jin, H., Leyva, A., Meisner, R.C.,  
Petroni, K., Urzainqui, A., Bevan, M., Martin, C., Smeekens, S.,  
Tonelli, C., Paz-Ares, J., and Weisshaar, B.  
Towards functional characterisation of the members of the R2R3-MYB  
gene family from Arabidopsis thaliana  
Plant J. 16 (2), 263-276 (1998)

JOURNAL  
MEDLINE  
99056848  
2 (bases 1 to 897)  
Meisner, R., Jin, H.-L., Martin, C. and Bevan, M.  
Direct Submission  
Submitted (04-May-1998) Abt. Biochemie, MPI fur Zuchtungsforchung,  
Carl-von-Linne-Weg 10, Koeln D-50829, Germany

FEATURES  
SOURCE Location/Qualifiers  
1.897  
/organism="Arabidopsis thaliana"  
/cultivar="Columbia"  
/db\_xref="taxon:3702"  
/chromosome="I"  
/map="near marker g3829"  
1.897  
/gene="MYB3"  
/allele="Col"  
17.757  
/gene="MYB3"  
/note="MYB3: R2R3-MYB factor family member"  
/codon\_start=1  
/product="putative transcription factor"  
/protein\_id="AAC83581.1"  
/db\_xref="GI:3941410"

gene  
1.897  
/gene="MYB3"  
/allele="Col"  
17.757  
/gene="MYB3"  
/note="MYB3: R2R3-MYB factor family member"  
/codon\_start=1  
/product="putative transcription factor"  
/protein\_id="AAC83581.1"  
/db\_xref="GI:3941410"

CDS  
/translation="MNGKAWTEEDERLVDYIRKHGCGMRSLPRAGLORCGKSCFL  
RNNYLRPLKGNFTEDDELITKLSLGNKWSLIAGLPERITN  
EIKWYNTHIRKRLINRGIDPTSHRPIQSSASQDSKPIOLEVTSMTINISFSAK  
VETFHESISFPGSEKISMLTFKEKDECPVOEKFDPDLNLEISLPDVRLOGHCK  
STPRCFKCSLGMINGMECRGMRCDVVGSSKSGSDMSGDFELGAKKETTSILGF  
DAVLCRCRIGLFNRNESCRCNCRVAVRTH"

BASE COUNT 246 a 153 c 235 g 263 t  
ORIGIN

Query Match 24.1% Score 242.4; DB 8; Length 897;  
Best Local Similarity 78.2% Pred. No. 7.1e-53;  
Matches 291; Conservative 0; Mismatches 81; Indels 0; Gaps 0;

QY 86 gctcattacaacaagagtcctgtgacaaagaggaagatacaagcctcattcaatac 145  
DB 11 GCTCATTACGAAAGAGAGCTTGAGACTTAAGAGAGAGAGAGAGAGAGATCA 70  
QY 146 cgtgtccatggtgagagtcgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 205  
DB 71 CGTAAACACGCTGAAGGTCCTGCGATCTCTCTCTCTCTCTCTCTCTCTCTCTCT 130  
QY 206 ggttaagttcagatgaatgaataaactactgtgagcctgtgattcttaagaagaaat 265  
DB 131 GGTAAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 190  
QY 266 ttaactgaagaagaatgagcttattcaatcaagcttcaagcttcaagcttcaagctt 325  
DB 191 TTAAGTGAAGAAGAAGATGATGATGATGATGATGATGATGATGATGATGATGAT 250  
QY 326 tcaattgattggtgaagatttaacaggaagacagataatgagataaagaactaggaac 385  
DB 251 TCAATTATAGCTGGAGATTAACAGGAAGACAGATTAACAGATTAACAGATTAAC 310  
QY 386 acacacataaagaagaagcttataagagagaaatgtgattccacaactatcgtctc 445  
DB 311 ACTCATATCAAGAGAGAGCTTCTCAGCCGTGGATGATGATGATGATGATGATGAT 370  
QY 446 aatcaaacggcc 457  
DB 371 AACGAATCCGTC 382

QY	415	aggaatgatccaaactcagcctcttcaa	447
DB	368	AGGATGTATCCAGCACCTCATAGAGGATCAA	400
RESULT	9		
LOCUS	AF034134	1151 bp	mrna
DEFINITION	Gossypium hirsutum MYB-like DNA-binding domain protein (Cmy-O)		
ACCESSION	AF034134		
VERSION	AF034134.1	GI:2921339	
KEYWORDS			
SOURCE	upland cotton.		
ORGANISM	Gossypium hirsutum		
REFERENCE	Eunaryota, Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II; Malvales; Malvaceae; Gossypium.		
AUTHORS	1 (bases 1 to 1151)		
TITLE	Loguercio, L.L., Zhang, J. and Wilkins, T.A.		
JOURNAL	Structure and expression of six classes of myb-domain genes in allotetraploid cotton (Gossypium hirsutum L.)		
REFERENCE	2 (bases 1 to 1151)		
AUTHORS	Loguercio, L.L., Zhang, J. and Wilkins, T.A.		
TITLE	Direct Submission		
JOURNAL	Submitted (13-NOV-1997) Agronomy & Range Science, University of California, One Shields Ave., Davis, CA 95616-8515, USA		
FEATURES	Location/Qualifiers		
source	1..1151		
gene	/organism="Gossypium hirsutum"		
CDS	/cultivar="Acacia, S1-2"		
	/db_xref="taxon:3635"		
	/dev_stage="3 days pre-anthesis"		
	/tissue_type="ovule"		
	1..1151		
	/gene="Cmy-O"		
	/note="MYB-domain gene O"		
	72..752		
	/gene="Cmy-O"		
	/function="putative MYB-like transcription factor"		
	/note="similar to MYB A encoded by GenBank Accession Number I04497"		
	/codon_start=1		
	/product="MYB-like DNA-binding domain protein"		
	/protein_id="AAC04720.1"		
	/db_xref="GI:2921340"		
	/translation="MRPKCCCKGCTNGANGSKQEDKLDYIRIHGEGCNSLPKRAAGLHRCGSKRLWYLPDRIKGNFAODEEDLILKLALGNRSLIAGRGSTDNENKVNWNSHKIRKLMKMGIDPNRHNKLNQYRPHVGPINPTTNSMDVACKLRCSNDNDGISAASYLEDATPPTGISINLDDLTITAFSPPIKNIIEESQCTASTIVNDEEOYTTVPTLPLR"		
	102..257		
	/note="putative MYB DNA-binding domain repeat R2"		
	258..412		
	/note="putative MYB DNA-binding domain repeat R3"		
BASE COUNT	362 a 215 c 240 g 314 t		
ORIGIN			
Query Match	21.6%	Score 217; DB 8; Length 1151;	
Best Local Similarity	71.3%	Pred. No. 3,36-46;	
Matches	286; Conservative	0; Mismatches 115; Indels 0; Gaps 0;	
QY	36	gattgatttgaatttgaactgattggaagatacactgttgytgaagaagcctatacca	95
DB	46	GTTTACTTTTGATTTGAAACTCCGAGAGAAAACCTTGCGATTAACAAGGACACA	105
QY	96	acaaagtgctgagccaaagagaagatcaagcctatacaactacactcgtgtccat	155
DB	106	ACAAAGGAGCGCTGTGCTCAAGCAAGAGATGCAAAACCTATTGATTATATACGATTATCAG	165

BASE COUNT	320	a	399	c	363	g	202	t
ORIGIN								
Query Match	21.5%	Score	216.4;	DB	8;	Length	1284;	
Best Local Similarity	71.8%;	Pred. No.	4.8e-46;					
Matches	283;	Conservative	0;	Mismatches	111;	Indels	0;	Gaps 0;
OY	54	aacgatggagcagctaccactgtgttgtaaaggccatacaacaagaagtgcctygacca	113					
Db	207	AAGCGATGGGAGGTGCCCTGCTCCGAGAAAGCCACACACAAGGCGCCTGGACCA	266					
OY	114	aagaagaatcaaaagcctcatcaactaatccgtytgcattgtaagtcyctggcgt	173					
Db	267	AGGAGAGAGACACGCAGCTCACCGCTACATCAAGGCGCAGCGCGGCTCTGGCGCT	326					
OY	174	cccccccgaagtcyctggcgcttagatgtgtgaagaagttgcagattaagatygata	233					
Db	327	CCCTGCCCAAGGCCCCCGGCTGCTCCGTCGGGGAAGGACGTCGCCCTCGATCA	386					
OY	234	actacttgaagcgttatctttaagagaagaaatctaactgaagaagaatgaagctatca	293					
Db	387	ACTACTCCGCCGCCACCTCAAAGCCGGCAATTCCGCCACAGAGAGAGAGCTCATCA	446					
OY	294	tcaagcttcacagttacttggtaaacaaatgttcattgattcttgyaagattaccagaa	353					
Db	447	TCAAGCTCCACAGCCTCCGCGGAACAATGTCCTTATACCGGGAGACATCGCGGGA	506					
OY	354	gaacagataatgatatgaagaactacttggacaacacacatcaagaagaagctataaga	413					
Db	507	GGACGGACACAAGATCAAGAATCTACTGGAAACGACACATCCGAGAGAACTGACAGACC	566					
OY	414	gaggaattgatccacaactcatgtccctctaa	447					
Db	567	GGGGATCGACCCGGTGACCCACCGCGCATCA	600					
RESULT	11							
LOCUS	D8617	1176 bp	MRNA	PLN	11-NOV-1997			
DEFINITION	Oryza sativa mRNA for OSMYB1, complete cds.							
ACCESSION	D8617							
VERSION	D8617.1 GI:2605616							
KEYWORDS	OSMYB1; OSMYB1.							
SOURCE	Oryza sativa seed cDNA to mRNA.							
ORGANISM	Oryza sativa							
REFERENCE								
AUTHORS	Suzuki,A.							
TITLE	Direct Submission							
JOURNAL	Submitted (28-OCT-1996) to the DDBJ/EMBL/GenBank databases. Akinhir Suzuki, National Institute of Agrobiological Resources, Department of Biotechnology, Kannondai 2-1-2, Tsukuba City 305, Japan (E-mail:hnaekht@agr.affrc.go.jp. Tel.:81-298-38-8450, Fax:81-298-38-8397)							
REFERENCE	2 (sites)							
AUTHORS	Suzuki,A., Suzukl,T., Tanabe,F., Toki,S., Washida,H., Wu,C.Y. and Takaiwa,F.							
TITLE	Cloning and expression of five myb-related genes from rice							
JOURNAL	Gene 198 (1-2), 393-398 (1997)							
FEATURES	98036074							
source	Location/Organisms 1..1176 /organism="Oryza sativa" /db_xref="taxon:4530"							

```
gene      /tissue_type="seed"  
          126. :845  
CDS       /gene="Osmyb1"  
          126. :845
```

```

function="transfactor"
/codon_start=1
/product="OSMyl1"
/protein_id="BAA23337.1"
/db_xref="GI:2605617"
/translation="MGRSCKGEAHTNKGAMWIKEDDRLIAYIRAHGEGCRLPKAN
GLRSGACSLRLMMVYLIRDLKICGAATDDEDEILRLSHLGLGKWSLIAGLPERTD
EIKANWNHKKRLALRGDIPDTRHRLVSGGDIASLNKHHRRRLIPSRRRRRP
SSPCAAALAAAFGRLLGRVLPQOQRHNEHGGAVPFPPALGRADAEALAGGDAHOR
IAGLELPLPRPRRGVQSLG"

```

BASE COUNT	253 a	357 c	355 g	211 t
ORIGIN				

Query Match	20.6%;	Score 206.8;	DB 8;	Length 1176;
Best Local Similarity	71.0%;	Pred. No. 1.5e-43;		
Matches 274;	Conservative	0;	Mismatches 112;	Indels 0;
			Gaps	0

OY	59	atggggaagatacaacccttggttgyaaaaaggtccataccaacaagaagtgccctgagccaagaag	111
Db	126	atgggggaaggtcccccctgcttgccgagaaaggccgacacagaaCAAGAGGGGGCCTGGACCAAGAG	185
OY	119	gaaga tcaagcgcctatccataactacatccgctgtccatgtgtgaagctgctgacgtccctc	178
Db	186	gagGACcAGcGGGCTATGCTACTACTCAAGGGCCGATGGCGAAGGCTGTGGCGCTCCCTG	245
OY	179	cccaaaagctgctgagcctgactagatgtgtgaagaattgcagattaaagatgataaactac	233
Db	246	CCCAAGGGGGGGGGGCTCTTCCTGCTGGCGCAAGAGCTGCCCTCCGGTGGATGAACTAC	305
OY	239	ttagagccttgatctttaagaagggaaatttcactcaagaagaagaatgagctatcaatgaag	288
Db	306	CTCCGCCCGCCGACTCTCAAGCGGGGCAACTCTCAACCGACGACGAGGACGGCTCATCATCCG	365
OY	299	cttcaacagtttacttggaaacaaatgtgtcatattgattctgtgaagatattccaggaagaagaca	358
Db	366	CTCCACAGCCCTCTCGGCAACAACTGGTCTGTATGCCCGGGCAGCTGCCGGGGAGGACG	425
OY	359	gataatagataaagaactactctggaacacacacatcaaaagaagaagctatataagcagagga	418
Db	426	GACACAGAGATCAAGAACTACTGTGACMACACGCACATCAAGCGCAAGGCTCTCGCCGGGGC	485
OY	419	atgatccaacaactcatgcttctct 444	
Db	486	ATGGACCCGGAGACGACCCGCCGCT 511	

RESULT	12
LOCUS	HVMYB2
DEFINITION	H. vulgare myb2 mRNA.
ACCESSION	X70876
VERSION	X70876.1 GI:19054
KEYWORDS	DNA-binding protein; myb DNA binding domain.
SOURCE	barley.
ORGANISM	Hordeum vulgare
REFERENCE	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; 1 (bases 1 to 1102)
AUTHORS	Rohde, W.
JOURNAL	Direct Submission
REFERENCE	Submitted (20-JAN-1993) W. Rohde, Max-Planck-Institut fuer Zuechtungsforschung, Carl-von-Linne-Weg 10, 5000 Koeln 30, FRG
AUTHORS	Rohde, W., Marocco, A., Wilsenbach, M., Barzen, E., Kristiansen, K. and Salami, F.
TITLE	Anthocyanin biosynthesis in barley: characterization of structural

JOURNAL	REFERENCE	AUTHORS	TITLE	JOURNAL	MEDLINE	COMMENT	FEATURES	SOURCE
(in) Styles, D. E., Gavazzi, G. and Racchi, M. S. (Eds.);	THE GENETICS OF FLAVONOIDS: 79-95;							
Edizioni Edicopli, Milano (1988)	3 (bases 1 to 1102)	Wissenbach, M.						
		Rohde, W.						
			Myb genes from Hordeum vulgare: tissue-specific expression of					
			chimeric Myb promoter/Gus genes in transgenic tobacco					
			Plant J. 4 (3), 411-422 (1993)					
			94035190					
			See X70876 and X70880 for related gene and cDNA sequences.					
			Location/Qualifiers					
			1. .1102					

BASE COUNT	224 a	356 c	350 g	172 t
ORIGIN				

Query Match	20.48;	Score 205.6;	DB 8;	Length 1102;
Best Local Similarity	70.68;	Pred. No. 3.2e-43;		
Matches 274;	Conservative 0;	Mismatches 114;	Indels 0;	Gaps 0;

QY	59	atgggaagcagtcaccccttggatgtgtaaaggctctatccaaacaaagtgctgtgacaaagag	11
Db	44	atggggaggtgccctctgctgcgaaagaaacacacacaaaggagggccttgacaccaaagag	10
QY	119	gaagatcaacgcctcatcaactacacaccggtgcacatgtgtgaagctgctggctgcctc	178
Db	104	gaggacacagcgcgctgattgcctacatcagggccaaagccgagggcctctggcgcgtccctg	168
QY	179	cccaaaagctgctgggctgctctagatgtgtgaaggttgacagattaaatgtatnaactac	238
Db	164	cccaagggggggggcctctctgcgcgcgacaaagactgcgccctccgctggatgaactac	223
QY	239	ttagggcttgatcctaaagagaggaattcactgaagaagaagatgtgattatcaatgaag	299
Db	224	ctccgcccccgaacctacacggcggaactctccacacacagcagggcagctcatcattccgc	283
QY	299	cttcacagtttacttgcgaacaaatggtcatgttatgtctgaagatattacacgaagaagaca	358
Db	284	ctccgacacgctgctgcggaacaaatggctcttattcccgccagctgcggggcagagacg	343
QY	359	gataatgtgataaagaactacttgcgaacacacacacatcaaaagaagagcttataagcagaagga	418
Db	344	gacacacgagatcaagaactacttgcgaacacgacacatcaagacgaagcagctctcccgccg	403
QY	419	attgatccaaactaatctgctccctctaa	446
Db	404	atggacccgacacgacacacggccctctca	431

RESULT 13

LOCUS D88618 1069 bp mRNA PLN 07-FEB-1999

DEFINITION Oryza sativa mRNA for OSMYB2, complete cds.

ACCESSION D88618

VERSION D88618.1 GI:2605618

KEYWORDS OsmYB2; OSMYB2.

SOURCE Oryza sativa seed cDNA to mRNA.

ORGANISM Oryza sativa

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Ehrhartoideae; Oryzaceae; Oryza.

1 (bases 1 to 1069)

REFERENCE

AUTHORS Suzuki, A.

TITLE Direct Submission

JOURNAL Submitted (28-OCT-1996) to the DDBJ/EMBL/GenBank databases. Akihiko Suzuki, National Institute of Agrobiological Resources, Department of Biotechnology, Kannondai 2-1-2, Tsukuba City 305, Japan (E-mail: hnsuk@abrr.affrc.go.jp, Tel: +81-298-38-8450, Fax: +81-298-38-8397)

2 (sites)

REFERENCE

AUTHORS Suzuki, T., Tanabe, F., Toki, S., Washida, H., Wu, C.Y. and Takaiwa, F.

TITLE Cloning and expression of five myb-related genes from rice seed

JOURNAL Gene 198 (1-2): 393-398 (1997)

FEATURES

source

1.1069 Location/Qualifiers

/organism="Oryza sativa"

/db\_xref="taxon:4530"

/rfeature="type=seed"

108..938

/gene="OsmYB2"

108..938

/gene="OsmYB2"

/function="transfactor"

/codon\_start=1

/product="OSMYB2"

/protein\_id="BA23338.1"

/db\_xref="GI:2605618"

/translation="MGRSPCCERKANTNKAWTEKEDORLIHINHGSGCWSLPKAA GILRGKSCRLRWNTNYLRDPLKRGNTFDEDELLITKLHLLGNKSLIAGRLPGRTD EIKRWNTHTKRLKLLAGLDPLRPARSMPPSPATTTSRRLQILARKNRKPPQQLR RRRVPRHRPQVPHKPAVVVPAVFAAARSGEIRSGEHLPPPLUGDESDM PVLEPPRAARRRVRVLRRLVGLFLLAAGDELASVHIROCIYVNSVKHREIQRSD"

BASE COUNT 230 a 375 c 289 g 175 t

ORIGIN

Query Match 20.0%; Score 201; DB 8; Length 1069;

Best Local Similarity 71.5%; Pred. No. 5,1e-42;

Matches 264; Conservative 0; Mismatches 105; Indels 0; Gaps 0;

58 gatggagcagcaccctgtgtgtgaaagcctcatccacaagaagctgtggaacaaaga 117

|||||

107 gatggagcagcaccctgtgtgtgaaagcctcatccacaagaagcctgtggaacaaaga 166

|||||

118 ggaagcagcagcctcatccactacatccgtgtcatgtggaagcgtgtggtggtccct 177

|||||

167 ggaagcagcagcctcatccactacatccgtgtcatgtggaagcgtgtggtggtccct 226

|||||

178 ccccaagctgtgtgtgtggtgtggtgtggtgtggtgtggtgtggtgtggtgtggtgtg 237

|||||

227 ccccaagcagcagcgt 286

|||||

238 cttagagcgt 297

|||||

287 ccttcgccccccacactcaagcagcagcagcagcagcagcagcagcagcagcagcagc 346

|||||

298 gcttcaagcttacttggaacaaatgtgtcatgtgtgtgtggaagattaccagaagaac 357

DB 347 GCTCCACAGACCTCCTCGGCAACAAGTGTGCTGATCGCCGGAGAGCTGCGGGAGAGC 406

QY 358 agaatatgagataaagactactctggaacacacacatcaaaaagactatagaagag 417

|||||

DB 407 GGACACAGAGATCAAGACTACTGGACACCCACATCAAGCCAAAGCTCCTCGCCCGCG 466

QY 418 aattgattcc 426

|||||

DB 467 CCTCGACCC 475

RESULT 14

ATMGENE 1335 bp mRNA PLN 17-MAR-2000

LOCUS Arabidopsis thaliana mRNA for putative transcription factor (MYB102 gene).

DEFINITION

ACCESSION X90382.2 GI:7263578

VERSION X90382.2

KEYWORDS myb-repeat; MYB102 gene; transcription factor.

SOURCE thale cress.

ORGANISM Arabidopsis thaliana

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis.

1 (bases 1 to 1335)

REFERENCE

AUTHORS Quaedvlieg, N., Dockx, J., Keultjes, G., Kock, P., Wilmering, J., Weisbeek, P. and Smeekens, S.

TITLE Identification of a light-regulated MYB gene from an Arabidopsis transcription factor gene collection

JOURNAL Plant Mol. Biol. 32 (5): 987-993 (1996)

97134989

2 (bases 1 to 1335)

REFERENCE

AUTHORS Keultjes, G.G.M.

TITLE Direct Submission

JOURNAL Submitted (28-JUL-1995) G.G.M. Keultjes, University of Utrecht, Padualaan 8, 3584 CH Utrecht, NETHERLANDS

3 (bases 1 to 1335)

REFERENCE

AUTHORS Smeekens, J.C.M.

TITLE Direct Submission

JOURNAL Submitted (17-MAR-2000) Smeekens J.C.M., Molecular Plant Physiology, Utrecht University, Padualaan 8, 3584 CH Utrecht, The Netherlands

On Mar 19, 2000 this sequence version replaced g1.928927.

COMMENT

FEATURES

source

1.1335 Location/Qualifiers

/organism="Arabidopsis thaliana"

/cultivar="Columbia"

/db\_xref="taxon:3702"

/chromosome="IV"

/note="allele Col"

join(1..241,242..371,372..1335)

/gene="MYB102"

1..1335

/gene="MYB102"

1..241

/gene="MYB102"

/number=1

join(109..241,242..371,372..1161)

/gene="MYB102"

/note="R2R3-MYB factor family member"

/codon\_start=1

/product="putative transcription factor"

/protein\_id="CAB81661.1"

/db\_xref="GI:7263579"

/translation="MARSPCCERKNGLKGPWTSEEDOKLVYIQRHGYNWRTLPKNA GIORGSGSRLRWNTNYLRDPLKRGNTFDEDELLITKLHLLGNKSLIAGRLPGRTD EIKRWNTHTKRLKLLAGLDPLRPARSMPPSPATTTSRRLQILARKNRKPPQQLR RRRVPRHRPQVPHKPAVVVPAVFAAARSGEIRSGEHLPPPLUGDESDM PVLEPPRAARRRVRVLRRLVGLFLLAAGDELASVHIROCIYVNSVKHREIQRSD MKFDIPFLDVNGFI"



XX 2



PF 01-DEC-2000; 2000MO-US32731.  
 XX 02-DEC-1999; 99US-0453387.  
 PR (REGC ) UNIV CALIFORNIA.  
 XX WILKINS TA;  
 XX WPI: 2001-397938/42.  
 DR P-PSDB; AAB82481.  
 XX  
 PT Modulating transcription in plants for improving cotton fiber quality,  
 PT comprises introducing a recombinant expression cassette comprising a  
 PT promoter operably linked to a sequence encoding a MYB polypeptide -  
 XX  
 XX Claim 3, Page 40-41; 50pp; English.  
 PS  
 CC The present sequence is that of cDNA encoding cotton putative MYB  
 CC transcription factor GhMYB 1 (see AAB82481). The cDNA was isolated  
 CC from a Gossypium hirsutum cv. Acala S2-3-day pre-anthesis ovule  
 CC cDNA library using a PCR-based strategy and Arabidopsis thaliana  
 CC Glabrous1 (AtMYB6/1) as a heterologous hybridisation probe. 6 MYB  
 CC genes, designated as GhMYB 1 to GhMYB 6, were identified. The  
 CC GhMYB 1, GhMYB 2 and GhMYB 3 genes are relatively abundant and  
 CC were expressed in all tissues examined; transcripts of the other  
 CC 3 genes are less abundant and exhibit tissue-specific patterns of  
 CC expression. A claimed method of modulating transcription in a  
 CC plant involves introducing into the plant a recombinant expression  
 CC cassette comprising a promoter operably linked to a heterologous  
 CC polynucleotide encoding a MYB protein, especially GhMYB 1 or  
 CC GhMYB 6. The plant is especially cotton, and the promoter directs  
 CC expression in cotton fibres or roots. Valuable phenotypes can  
 CC be conferred on the plant by this method, e.g. increased fibre  
 CC quality, alteration of root architecture, and enhanced growth.  
 CC The GhMYB1 gene is particularly useful for increasing levels of  
 CC gene expression in plant fibres to increase fibre yield, length,  
 CC strength and fineness.  
 CC  
 SO Sequence 1006 BP; 323 A; 209 C; 204 G; 270 T; 0 other;

Query Match 100.0%; Score 1006; DB 22; Length 1006;  
 Best Local Similarity 100.0%; Pred. No. 5 4e-271;  
 Matches 1006; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 taacacggtatctctctatcttctacgtgattgttgattgttaactgat 60  
 DB 1 taacacggtatctctctatcttctacgtgattgttgattgttaactgat 60  
 OY 61 gggagatctactctgttggaaaggtctatcaacaagaagtgccggacaagaaga 120  
 DB 61 gggagatctactctgttggaaaggtctatcaacaagaagtgccggacaagaaga 120  
 OY 121 agatcaacgctcatcaactacatccgtgcatgtgtaagggcgtgcgtccctccc 180  
 DB 121 agatcaacgctcatcaactacatccgtgcatgtgtaagggcgtgcgtccctccc 180  
 OY 181 caaagctgtcggcgtccttagatgtgtgaagagttgcagattaagatgaataactact 240  
 DB 181 caaagctgtcggcgtccttagatgtgtgaagagttgcagattaagatgaataactact 240  
 OY 241 gaggcctgacttaagaagaggaattcactgaagaagaagatgagcttaccatcaact 300  
 DB 241 gaggcctgacttaagaagaggaattcactgaagaagaagatgagcttaccatcaact 300  
 OY 301 tcacagttacttggaaacaatggtcatgtgctggaagattaccaggaagaacaga 360  
 DB 301 tcacagttacttggaaacaatggtcatgtgctggaagattaccaggaagaacaga 360  
 OY 361 taatgagataaagaactactcgtgaacacacacatacaagaagaagcttataagcagaagaat 420  
 DB 361 taatgagataaagaactactcgtgaacacacacatacaagaagaagcttataagcagaagaat 420

OY 421 tgatccacaactacatcgtctctctcaatcaaacgycgaataccacaacagtcacagcccc 480  
 DB 421 tgatccacaactacatcgtctctctcaatcaaacgycgaataccacaacagtcacagcccc 480  
 OY 481 caacgaattggtattcaagaactcgcgcacatccggttccaaatccagttccatcaaaaa 540  
 DB 481 caacgaattggtattcaagaactcgcgcacatccggttccaaatccagttccatcaaaaa 540  
 OY 541 ccggtctcgtgatttcaatataatgatttcaattcaatgaatccaaacacagttcccttga 600  
 DB 541 ccggtctcgtgatttcaatataatgatttcaattcaatgaatccaaacacagttcccttga 600  
 OY 601 agaacccaactgacagccagcagatgacatgactacagatgaagaagcacaagaacagct 660  
 DB 601 agaacccaactgacagccagcagatgacatgactacagatgaagaagcacaagaacagct 660  
 OY 661 gcaacaagaagcagcaatagcgtccgagcaatgaggcaagaacataattgtagctgcat 720  
 DB 661 gcaacaagaagcagcaatagcgtccgagcaatgaggcaagaacataattgtagctgcat 720  
 OY 721 tgggattgttcagctgactcactcgcggtatcaaatgccaactcgcgcgagtcgaacc 780  
 DB 721 tgggattgttcagctgactcactcgcggtatcaaatgccaactcgcgcgagtcgaacc 780  
 OY 781 aaagtagataacaacaatttccagttcttgaacaaagctatggtggtgaagcggtatg 840  
 DB 781 aaagtagataacaacaatttccagttcttgaacaaagctatggtggtgaagcggtatg 840  
 OY 841 ttgtgttgcaattaggttttggaaacagtgaaattgttagaactgtccaaattcaaa 900  
 DB 841 ttgtgttgcaattaggttttggaaacagtgaaattgttagaactgtccaaattcaaa 900  
 OY 901 tccaatgctctcatagttatgttagaaccttgatcatcagggtcattcttctct 960  
 DB 901 tccaatgctctcatagttatgttagaaccttgatcatcagggtcattcttctct 960  
 OY 961 tctcttctgttttagagataaattatgtcttaattatataaaaaa 1006  
 DB 961 tctcttctgttttagagataaattatgtcttaattatataaaaaa 1006

RESULT 2  
 AAF80412  
 ID AAF80412 standard; cDNA; 764 BP.

XX AAF80412;  
 XX  
 XX 29-JUN-2001 (first entry)  
 DE Nucleotide sequence of a plant transcription factor G208:  
 XX  
 XX Transcription factor; flowering time; transgenic plant; vernalisation;  
 KW plant development; plant physiology; flowering; ss.  
 OS Arabidopsis thaliana.  
 FH  
 FT Key Location/Qualifiers  
 FT CDS 15..725  
 FT /tag= a  
 FT /product= "transcription factor"  
 PN  
 XX  
 XX MO200126459-A2.  
 PD 19-APR-2001.  
 XX  
 XX  
 PF 11-OCT-2000; 2000MO-US28141.  
 XX  
 XX 12-OCT-1999; 99US-0159464.  
 PR 08-NOV-1999; 99US-0164132.  
 PR 17-NOV-1999; 99US-0166228.  
 PR 17-APR-2000; 2000US-0197899.  
 PR 22-AUG-2000; 2000US-0227439.  
 XX

Thu Jan 24 07:39:28 2002

us-09-453-387a-1.rng

Page 3

PA (MEND-) MENDEL BIOTECHNOLOGY INC.  
 PA (RATC/) RATCLIFFE O.  
 PA (HEAR/) HEARD J.  
 PA (SAMA/) SAMAHA R.  
 PA (CREE/) CREELMAN R.  
 PA (KEDD/) KEDDIE J.  
 PA (JIANG/) JIANG C.  
 PA (REUB/) REUBER L.  
 PA (RIEC/) RIECHMANN J L.  
 XX  
 PI Ratcliffe O, Heard J, Samaha R, Creelman R, Keddie J, Jiang C;  
 PI Reuber L, Riechmann JL;  
 XX  
 DR WPI, 2001-266398/27.  
 XX P-PSDB; AAB67842.  
 XX  
 PT New transgenic plant comprises a recombinant polynucleotide encoding a  
 PT plant transcription factor polypeptide and has a modified flowering  
 PT time or vernalization requirement  
 XX  
 PS Example 7; Page 104-105; 108pp; English.  
 XX  
 CC The present sequence encodes a plant transcription factor protein  
 CC which modifies the flowering time of a plant. The polynucleotide  
 CC sequence is used to produce transgenic plants which have a modified  
 CC flowering time or a modified vernalization requirement. The  
 CC polynucleotides and polypeptides are useful for modifying plant  
 CC development, physiology or biochemistry such that the modified plants  
 CC have a trait advantage over wild type plants. In particular they are  
 CC useful for accelerating, delaying or preventing flowering. The  
 CC polynucleotides are also useful as nucleic acid probes and primers.  
 CC They may be used to identify proteins that can modify the activity of  
 CC the transcription factor.  
 XX  
 SQ Sequence 764 BP; 228 A; 140 C; 177 G; 219 T; 0 other;

Query Match 26.2%; Score 263.8; DB 22; Length 764;  
 Best Local Similarity 79.2%; Pred. No. 8e-64;  
 Matches 313; Conservative 0; Mismatches 82; Indels 0; Gaps 0;

QY 59 atggagcagctacacctgtgtgaaagcctacacacaaagctgctgagccaaagag 118  
 DB 15 atggagagctcctcctgtgtgaaagcctacacaaagagctgagcctaagaa 74  
 QY 119 gaagatcaagcctcaacatacatcgtgtccatgtgaaagctgctgctgctcc 178  
 DB 75 gaagatcaagcctcgtagattatccgttaacacggtgaaagctgtgtgcttct 134  
 QY 179 ccgaagcctgctgagctagatgtgaaagctgcaagtaagatgagtaaacac 238  
 DB 135 cctaagcctgctgagctgtgtgaaagctgcaagtaagatgagtaaacac 194  
 QY 239 ttgagcctgctgcttaagagagaattcactgaagagaagatgagcttaacaa 298  
 DB 195 ctctgcctgctgcttaagagagaattcactgagtaagatgagtaaacac 254  
 QY 299 ctccacagcttacttgaagagaattcactgagtaagatgagtaaacac 358  
 DB 255 ctccacagcttacttgaagagaattcactgagtaagatgagtaaacac 314  
 QY 359 gataatgaatgaagagaattcactgagtaagatgagtaaacac 418  
 DB 315 gataatgaatgaagagaattcactgagtaagatgagtaaacac 374  
 QY 419 atgatccacaactcactgctcctcaatcaac 453  
 DB 375 atgatccacaactcactgctcctcaatcaac 409

RESULT 3  
 AAC49989  
 ID AAC49989 standard; DNA; 868 BP.

XX AAC49989;  
 AC 18-OCT-2000 (first entry)  
 XX  
 DT Arabidopsis thaliana DNA fragment SEQ ID NO: 63187.  
 XX  
 DE Arabidopsis thaliana DNA fragment SEQ ID NO: 63187.  
 XX  
 KW Hybridisation assay; genetic mapping; gene expression control;  
 KW protein identification; signal transduction pathway;  
 KW metabolic pathway; promoter; termination sequence; ss.  
 XX  
 OS Arabidopsis thaliana.  
 XX  
 PN EP1033405-A2.  
 XX  
 PD 06-SEP-2000.  
 XX  
 PE 25-FEB-2000; 2000EP-0301439.  
 XX  
 XX 25-FEB-1999; 99US-0121825.  
 PR 05-MAR-1999; 99US-0123180.  
 PR 09-MAR-1999; 99US-0123548.  
 PR 23-MAR-1999; 99US-0125788.  
 PR 25-MAR-1999; 99US-0126284.  
 PR 29-MAR-1999; 99US-0126785.  
 PR 01-APR-1999; 99US-0127462.  
 PR 08-APR-1999; 99US-0128234.  
 PR 15-APR-1999; 99US-0128714.  
 PR 16-APR-1999; 99US-0129845.  
 PR 19-APR-1999; 99US-0130077.  
 PR 21-APR-1999; 99US-0130449.  
 PR 23-APR-1999; 99US-0130510.  
 PR 23-APR-1999; 99US-0130891.  
 PR 28-APR-1999; 99US-0131449.  
 PR 30-APR-1999; 99US-0132048.  
 PR 30-APR-1999; 99US-0132407.  
 PR 04-MAY-1999; 99US-0132484.  
 PR 05-MAY-1999; 99US-0132485.  
 PR 06-MAY-1999; 99US-0132486.  
 PR 07-MAY-1999; 99US-0132487.  
 PR 07-MAY-1999; 99US-0132863.  
 PR 11-MAY-1999; 99US-0134256.  
 PR 14-MAY-1999; 99US-0134218.  
 PR 14-MAY-1999; 99US-0134219.  
 PR 14-MAY-1999; 99US-0134221.  
 PR 14-MAY-1999; 99US-0134370.  
 PR 18-MAY-1999; 99US-0134768.  
 PR 19-MAY-1999; 99US-0134941.  
 PR 20-MAY-1999; 99US-0135124.  
 PR 21-MAY-1999; 99US-0135353.  
 PR 24-MAY-1999; 99US-0135629.  
 PR 25-MAY-1999; 99US-0136021.  
 PR 27-MAY-1999; 99US-0136382.  
 PR 28-MAY-1999; 99US-0136782.  
 PR 01-JUN-1999; 99US-0137222.  
 PR 03-JUN-1999; 99US-0137528.  
 PR 04-JUN-1999; 99US-0137502.  
 PR 07-JUN-1999; 99US-0137724.  
 PR 08-JUN-1999; 99US-0138094.  
 PR 10-JUN-1999; 99US-0138540.  
 PR 10-JUN-1999; 99US-0138847.  
 PR 14-JUN-1999; 99US-0139119.  
 PR 16-JUN-1999; 99US-0139452.  
 PR 16-JUN-1999; 99US-0139453.  
 PR 17-JUN-1999; 99US-0139452.  
 PR 18-JUN-1999; 99US-0139454.  
 PR 18-JUN-1999; 99US-0139455.  
 PR 18-JUN-1999; 99US-0139456.  
 PR 18-JUN-1999; 99US-0139457.  
 PR 18-JUN-1999; 99US-0139458.  
 PR 18-JUN-1999; 99US-0139459.  
 PR 18-JUN-1999; 99US-0139460.  
 PR 18-JUN-1999; 99US-0139461.

PR 18-JUN-1999; 99US-0139462.  
 PR 18-JUN-1999; 99US-0139463.  
 PR 18-JUN-1999; 99US-0139750.  
 PR 18-JUN-1999; 99US-0139763.  
 PR 21-JUN-1999; 99US-0139817.  
 PR 22-JUN-1999; 99US-0139899.  
 PR 23-JUN-1999; 99US-0140353.  
 PR 23-JUN-1999; 99US-0140354.  
 PR 24-JUN-1999; 99US-0140695.  
 PR 28-JUN-1999; 99US-0140823.  
 PR 29-JUN-1999; 99US-0140991.  
 PR 30-JUN-1999; 99US-0141287.  
 PR 01-JUL-1999; 99US-0141842.  
 PR 01-JUL-1999; 99US-0142154.  
 PR 02-JUL-1999; 99US-0142055.  
 PR 06-JUL-1999; 99US-0142390.  
 PR 08-JUL-1999; 99US-0142803.  
 PR 09-JUL-1999; 99US-0142920.  
 PR 12-JUL-1999; 99US-0142977.  
 PR 13-JUL-1999; 99US-0143542.  
 PR 14-JUL-1999; 99US-0143624.  
 PR 15-JUL-1999; 99US-0144005.  
 PR 16-JUL-1999; 99US-0144085.  
 PR 16-JUL-1999; 99US-0144086.  
 PR 19-JUL-1999; 99US-0144325.  
 PR 19-JUL-1999; 99US-0144331.  
 PR 19-JUL-1999; 99US-0144332.  
 PR 19-JUL-1999; 99US-0144333.  
 PR 19-JUL-1999; 99US-0144334.  
 PR 19-JUL-1999; 99US-0144335.  
 PR 20-JUL-1999; 99US-0144352.  
 PR 20-JUL-1999; 99US-0144632.  
 PR 20-JUL-1999; 99US-0144684.  
 PR 21-JUL-1999; 99US-0144814.  
 PR 21-JUL-1999; 99US-0145086.  
 PR 21-JUL-1999; 99US-0145088.  
 PR 22-JUL-1999; 99US-0145085.  
 PR 22-JUL-1999; 99US-0145087.  
 PR 22-JUL-1999; 99US-0145089.  
 PR 22-JUL-1999; 99US-0145192.  
 PR 23-JUL-1999; 99US-0145145.  
 PR 23-JUL-1999; 99US-0145218.  
 PR 26-JUL-1999; 99US-0145224.  
 PR 26-JUL-1999; 99US-0145276.  
 PR 27-JUL-1999; 99US-0145913.  
 PR 27-JUL-1999; 99US-0145918.  
 PR 27-JUL-1999; 99US-0145919.  
 PR 28-JUL-1999; 99US-0145951.  
 PR 02-AUG-1999; 99US-0146386.  
 PR 02-AUG-1999; 99US-0146388.  
 PR 03-AUG-1999; 99US-0146389.  
 PR 03-AUG-1999; 99US-0147038.  
 PR 04-AUG-1999; 99US-0147204.  
 PR 04-AUG-1999; 99US-0147302.  
 PR 05-AUG-1999; 99US-0147192.  
 PR 05-AUG-1999; 99US-0147260.  
 PR 06-AUG-1999; 99US-0147303.  
 PR 06-AUG-1999; 99US-0147416.  
 PR 09-AUG-1999; 99US-0147493.  
 PR 09-AUG-1999; 99US-0147935.  
 PR 10-AUG-1999; 99US-0148171.  
 PR 11-AUG-1999; 99US-0148319.  
 PR 12-AUG-1999; 99US-0148341.  
 PR 13-AUG-1999; 99US-0148565.  
 PR 13-AUG-1999; 99US-0148684.  
 PR 16-AUG-1999; 99US-0149368.  
 PR 17-AUG-1999; 99US-0149175.  
 PR 18-AUG-1999; 99US-0149426.  
 PR 20-AUG-1999; 99US-0149722.  
 PR 20-AUG-1999; 99US-0149723.  
 PR 20-AUG-1999; 99US-0149929.  
 PR 23-AUG-1999; 99US-0149902.  
 PR 23-AUG-1999; 99US-0149930.

PR 25-AUG-1999; 99US-0150566.  
 PR 26-AUG-1999; 99US-0150884.  
 PR 27-AUG-1999; 99US-0151065.  
 PR 27-AUG-1999; 99US-0151066.  
 PR 27-AUG-1999; 99US-0151080.  
 PR 30-AUG-1999; 99US-0151303.  
 PR 31-AUG-1999; 99US-0151338.  
 PR 01-SEP-1999; 99US-0151930.  
 PR 07-SEP-1999; 99US-0152363.  
 PR 10-SEP-1999; 99US-0153070.  
 PR 13-SEP-1999; 99US-0153758.  
 PR 15-SEP-1999; 99US-0154018.  
 PR 16-SEP-1999; 99US-0154039.  
 PR 20-SEP-1999; 99US-0154779.  
 PR 22-SEP-1999; 99US-0155139.  
 PR 23-SEP-1999; 99US-0155486.  
 PR 24-SEP-1999; 99US-0155659.  
 PR 28-SEP-1999; 99US-0156458.  
 PR 29-SEP-1999; 99US-0156596.  
 PR 04-OCT-1999; 99US-0157117.  
 PR 05-OCT-1999; 99US-0157753.  
 PR 06-OCT-1999; 99US-0157865.  
 PR 07-OCT-1999; 99US-0158029.  
 PR 08-OCT-1999; 99US-0158232.  
 PR 12-OCT-1999; 99US-0158369.  
 PR 13-OCT-1999; 99US-0159293.  
 PR 13-OCT-1999; 99US-0159294.  
 PR 14-OCT-1999; 99US-0159295.  
 PR 14-OCT-1999; 99US-0159329.  
 PR 14-OCT-1999; 99US-0159330.  
 PR 14-OCT-1999; 99US-0159331.  
 PR 14-OCT-1999; 99US-0159637.  
 PR 14-OCT-1999; 99US-0159638.  
 PR 18-OCT-1999; 99US-0159584.  
 PR 21-OCT-1999; 99US-0160741.  
 PR 21-OCT-1999; 99US-0160767.  
 PR 21-OCT-1999; 99US-0160768.  
 PR 21-OCT-1999; 99US-0160770.  
 PR 21-OCT-1999; 99US-0160814.  
 PR 21-OCT-1999; 99US-0160815.  
 PR 21-OCT-1999; 99US-0160816.  
 PR 22-OCT-1999; 99US-0160981.  
 PR 22-OCT-1999; 99US-0160989.  
 PR 22-OCT-1999; 99US-0161404.  
 PR 25-OCT-1999; 99US-0161405.  
 PR 25-OCT-1999; 99US-0161406.  
 PR 26-OCT-1999; 99US-0161359.  
 PR 26-OCT-1999; 99US-0161360.  
 PR 26-OCT-1999; 99US-0161361.  
 PR 28-OCT-1999; 99US-0161920.  
 PR 28-OCT-1999; 99US-0161992.  
 PR 28-OCT-1999; 99US-0161993.  
 PR 28-OCT-1999; 99US-0162142.

Query Match 26.2%; Score 263.8; DB 21; Length 866;  
 Best Local Similarity 79.2%; Pred. No. 8.5e-64;  
 Matches 313; Conservative 0; Mismatches 82; Indels 0; Gaps 0;

QY 59 atggagcagatcactctgtgtgaaaggctcctaccacaagaagctgagcccaagag 118  
 Db 52 atggagagatctctctgtgtgaaaggctcaccacaagaagcttggactaagaa 111  
 QY 119 gaagatcaacgcctcatcactacatccgtgtccatgtgtgaaggctgctgctccctc 178  
 Db 112 gaagatcaacgcctcgtatgatattatccgtatcaacgcgtgaaggtgtgctctctc 171  
 QY 179 cccaagctgctgctgctgaatgagatgtgtgaagagttgcagattgaagatgaactac 236  
 Db 172 cctaactcgcgtgattgtgtgtgtgaagattgtgagattgagattgaattac 231  
 QY 239 ttgagcctgattcttaagaaggaaattcactgagaagaagaagatgattcatcaag 298  
 Db 232 ctgcctcgtatcttaacggtgtgaatttactgattgattgaagatcaatcatcaaaa 291



	Matches	313;	Conservative	0;	Mismatches	82;	Indels	0;	Gaps
OY	59	atgagcgcatacctctgttctggaagaagctcataccacaagaagtgcctgagccaagaag	118						
Db	54	atgggaagatctctctgttctggaagaagctcataccacaagaagaagcttgaactaaga	113						
OY	119	gaagatcaacgcctcatcaactcatccgtgcgtccatggtgaaggtcgtgcttccctc	178						
Db	114	gaagatcaacgctctctgtagattatctccgaatcaccggtgaaggtctgtgcttccctc	173						
OY	179	cccaaaagctctggtgcgcgttagatgtgtgaagaagttgcagatcaaatgatcaaac	238						
Db	174	cccaaaatccgcctgtgattgttgcgttctgtggaacaaagttgattgagatgaattac	233						
OY	239	ttgagcctcatctttaagagagaaatttctacttgaagaagaatgagcttcatcaatga	298						
Db	234	cttcgtctccatcttctaagaagctgtggaatttctcatgatgatgaagctcaatcatca	293						
OY	299	cttcacagtttactcttggaaacaaatgtcatattgtctctggaagattaccagaaga	358						
Db	294	ctccatagcttactctgtgaacaaatgtctcatlgtatgtggaagattaccagaaga	353						
OY	359	gataatagatataagaactactctggaacacacacgtcaaaagaagcttataagaaga	418						
Db	354	gataacgaataaagaattatttgaacacacccatcatataagaagaagcttcttagtca	413						
OY	419	atgatccacaactcatcgtccctcatcaatcaaac	453						
Db	414	atgatccacaactcatcgtccctcatcaatcaaac	448						
RESULT	5								
ID	AAAC57347								
ID	AAAC57347	standard; DNA: 473 BP.							
AC	AAAC57347;								
DT	25-JAN-2001	(first entry)							
XX									
DE		Pinus radiata transcription factor DNA sequence #702.							
KM		Plant; transcription factor; gene expression; eucalyptus; pine; acacia;							
KM		poplar; sweetgum; teak; mahogany; bZIP; G-box binding factor;							
KM		basic helix-loop-helix zipper; homeotic; homeodomain; homeobox; MADS;							
KM		homeodomain zipper; LIM domain; AP2; ERBs; zinc finger domain;							
KM		type 2 Cys2His2; CCAAT box element; MYB; ss.							
OS		Pinus radiata.							
XX									
XX		MO200053724-A2.							
PD									
XX		14-SEP-2000.							
XX									
XX		09-MAR-2000; 2000MO-US06112.							
XX									
XX		11-MAR-1999; 99US-0266513.							
PR		18-AUG-1999; 99US-0149485.							
XX									
PA		(GENE-) GENESIS RES & DEV CORP LTD.							
PA		(FLET-) FLETCHER CHALLENGE FORESTS LTD.							
XX									
PI		Wood M, McGrath A., Shenk MA, Glenn M;							
XX									
DR		WPI: 2000-579369/54.							
XX									
PT		New isolated polynucleotide encoding a plant transcription factor for							
PT		producing a plant e.g. a woody plant, preferably eucalyptus or pine,							
PT		having modified gene expression or modified activity of a polypeptide							
XX									
XX									
XX									
CC		Claim 1; Page 653; 747pp; English.							
CC		The present invention relates to novel plant transcription factors from							

Sequence 473 BP; 131 A; 134 C; 120 G; 88 T; 0 other;

OY 54 aacacgataaggagacatcaacacctgtgtgtgaaaggctctataccaacaaagtgcttgagcca 113  
 DB 35 aaaaagatagggagatctctccgtgtgtgtgaaaggctctataccaacaaagggccttgagcta 94  
 OY 114 aagagagagatcaacgctctatacaactacatcgtgtgtgtgtgaaggtgtgtgtgcgtt 173  
 DB 95 aacaaagaaagacgacgcgtcttatcgctcacatctgaagccacggcgaaggggcttcgcgtt 154  
 OY 174 ccccccacaagctgcgtgcgtgtctttagatgtgtgtaaaggttcagataagatagatgataa 233  
 DB 155 cgtctcccaagtcgcgcagggctgtcttgaaatgctgcgaagagctgtgcagctgcgttgatataa 214

Db	215	actatccgcgtccgcgactctgaagcgtggaagcttcaacggagaagaagagctcatca	274
OY	234	actctctggaagccctgagactcttaagaagagaaattctactgtgaagaagaagatgaattaca	293
Db	215		274
OY	294	tcaagcttcacagttctattcttggaacaacaaatgvtcatatgtatctgtggaagattccaggaa	353
Db	275		334
OY	354	gaacagatgaatgagataaagaataactctggaacacacacatcacaagaagcttataagca	413
Db	335	ggaacgagcaacagagataaagaacatcaactctggaacacacacatcacaagaagaatctctgagca	394
OY	414	gaggaaattgtatccacaacatccatctgctctctatcaatacaagaagcccatacc	463
Db	395	agggagactcgagaccaccaacccatctgtctcatctggtcgagcccaaacatacc	444

	RESULT	6
XX	AAC57075	
ID	AAC57075	standard; DNA; 509 BP.
XX		
AC	AAC57075;	
XX		
DT	25-JAN-2001	(first entry)
XX		
DE	Plnus radiata transcription factor DNA sequence #521.	
XX		
KW	Plant; transcription factor; gene expression; eucalyptus; plne; acacia;	
KW	poplar; sweetgum; teak; mahogany; bzrp; G-box binding factor;	
KW	basic helix-loop-helix zipper; homeotic; homeodomain; homeobox; MADS;	
KW	homeodomain zipper; Lim domain; AP2; EREBS; zinc finger domain;	
KW	type 2 Cys2His2; CCAAT box element; MYB; ss.	
XX		
OS	Plnus radiata.	
XX		
PN	WO200053724-A2.	
XX		
PD	14-SEP-2000.	
XX		
PF	09-MAR-2000; 2000WO-US0612.	
XX		
PR	11-MAR-1999; 99US-0266513.	

DR WPI; 2000-579369/54.

PS claim 1: Page 580: 747pp: English

CC The present invention relates to novel plant transcription factors from  
CC *Eucalyptus grandis* or *Pinus radiata*. The present sequence is the coding  
CC sequence for one such transcription factor. The transcription factor may  
CC be used to produce a plant having modified gene expression such as a  
CC woody plant e.g. a *eucalyptus*, pine, acacia, poplar, sweetgum, teak, or  
CC mahogany species or to modify the activity of a polypeptide in a plant.  
CC The transcription factors of the present invention are members from the  
CC following families of regulatory proteins: bZIP, bZIP family of G-box  
CC binding factors, basic helix-loop-helix zipper,  
CC homeotic/homeodomain/homeobox/MADS, homeodomain zipper, LIM domain, AP2  
CC and E2Fs, zinc finger domains of type 2 Cys2His2, CCAAT box elements  
CC and MYB.  
XX  
XX  
50 Sequence 509 BP; 141 A; 147 C; 125 G; 96 T; 0 other;

Query Match	26.0%	Score 261.2;	DB 21;	Length 509;
Best Local Similarity	77.3%	Pred. No. 3.6e-63;		
Matches 317; Conservative	0;	Mismatches 99;	Indels 0;	Gaps 0

[illegible]

RESULT	7
AAC57365	
ID	AAC57365 standard; DNA; 1576 BP
XX	
AC	
AAC57365;	
XX	
DT	
XX	
XX	25-JAN-2001 (first entry)
XX	

DE Pinus radiata transcription factor DNA sequence #719.  
 XX  
 XX Plant; transcription factor; gene expression; eucalyptus; pine; acacia;  
 KW poplar; sweetgum; teak; mahogany; bZIP; G-box binding factor;  
 KW basic helix-loop-helix zipper; homeotic; homeobox; MADS;  
 KW homeodomain zipper; LIM domain; AP2; EREBS; zinc finger domain;  
 KW type 2 Cys2His2; CCAAT box element; MYB; ss.  
 XX  
 OS Pinus radiata.  
 OS  
 PN W0200053724-A2.  
 XX  
 XX 14-SEP-2000.  
 PD  
 PE 09-MAR-2000; 2000MO-US06112.  
 XX  
 XX 11-MAR-1999; 99US-0266513.  
 PR 18-AUG-1999; 99US-0149485.  
 XX  
 PA (GENE-) GENESIS RES & DEV CORP LTD.  
 PA (FLET-) FLETCHER CHALLENGE FORESTS LTD.  
 XX  
 PI Wood M, McGrath A, Shenk MA, Glenn M;  
 DR WPI; 2000-579369/54.  
 XX  
 XX New isolated polynucleotide encoding a plant transcription factor for  
 PT producing a plant e.g. a woody plant, preferably eucalyptus or pine,  
 PT having modified gene expression or modified activity of a polypeptide  
 PT  
 PS  
 XX  
 XX Claim 1; Page 658; 747pp; English.  
 CC The present invention relates to novel plant transcription factors from  
 CC Eucalyptus grandis or Pinus radiata. The present sequence is the coding  
 CC sequence for one such transcription factor. The transcription factor may  
 CC be used to produce a plant having modified gene expression such as a  
 CC woody plant e.g. a eucalyptus, pine, acacia, poplar, sweetgum, teak, or  
 CC mahogany species or to modify the activity of a polypeptide in a plant.  
 CC The transcription factors of the present invention are members from the  
 CC following families of regulatory proteins: bZIP, bZIP family of G-box  
 CC binding factors, basic helix-loop-helix zipper,  
 CC homeotic/homeodomain/homeobox/MADS, homeodomain zipper, LIM domain, AP2  
 CC and EREBS, zinc finger domains of type 2 Cys2His2, CCAAT box elements  
 CC and MYB.  
 CC  
 XX  
 XX Sequence 1576 BP; 456 A; 349 C; 382 G; 389 T; 0 other;  
 SQ

Query Match 26.0%; Score 261.2; DB 21; Length 1576;  
 Best Local Similarity 77.3%; Pred. No. 5.9e-63;  
 Matches 317; Conservative 0; Mismatches 93; Indels 0; Gaps 0;

QY 354 gaacagataatgataaagaactctggaacacacacatcaaaagaactataagca 413  
 Db 335 gagagacaacagataaagaactactggaacacacacatcaaaagaactgctgagca 394  
 QY 414 gagagatgatacacaactatcgtccctcctcaatcaaaagggccaatcc 463  
 Db 395 agggactcgaccaccaaccatcgtccactcgtcgccagcaagaacaataacc 444  
 RESULT 8  
 AAC42662  
 ID AAC42662 standard; DNA; 774 BP.  
 XX  
 AC AAC42662;  
 XX  
 XX 17-OCT-2000 (first entry)  
 DE Arabidopsis thaliana DNA fragment SEQ ID NO: 36392.  
 DE  
 KW Hybridisation assay; genetic mapping; gene expression control;  
 KW protein identification; signal transduction pathway;  
 KW metabolic pathway; promoter; termination sequence; ss.  
 XX  
 XX Arabidopsis thaliana.  
 OS  
 PN EP1033405-A2.  
 XX  
 XX 06-SEP-2000.  
 PD  
 PE 25-FEB-2000; 2000EP-0301439.  
 XX  
 XX 25-FEB-1999; 99US-0121825.  
 PR 05-MAR-1999; 99US-0123180.  
 PR 09-MAR-1999; 99US-0123548.  
 PR 23-MAR-1999; 99US-0125788.  
 PR 25-MAR-1999; 99US-0126264.  
 PR 29-MAR-1999; 99US-0126785.  
 PR 01-APR-1999; 99US-0127462.  
 PR 06-APR-1999; 99US-0128234.  
 PR 08-APR-1999; 99US-0128714.  
 PR 16-APR-1999; 99US-0129845.  
 PR 19-APR-1999; 99US-0130077.  
 PR 21-APR-1999; 99US-0130449.  
 PR 23-APR-1999; 99US-0130510.  
 PR 23-APR-1999; 99US-0130891.  
 PR 28-APR-1999; 99US-0131449.  
 PR 30-APR-1999; 99US-0132048.  
 PR 30-APR-1999; 99US-0132407.  
 PR 04-MAY-1999; 99US-0132484.  
 PR 05-MAY-1999; 99US-0132485.  
 PR 06-MAY-1999; 99US-0132486.  
 PR 06-MAY-1999; 99US-0132487.  
 PR 07-MAY-1999; 99US-0132863.  
 PR 11-MAY-1999; 99US-0134256.  
 PR 14-MAY-1999; 99US-0134218.  
 PR 14-MAY-1999; 99US-0134219.  
 PR 14-MAY-1999; 99US-0134221.  
 PR 14-MAY-1999; 99US-0134370.  
 PR 18-MAY-1999; 99US-0134768.  
 PR 19-MAY-1999; 99US-0134941.  
 PR 20-MAY-1999; 99US-0135124.  
 PR 21-MAY-1999; 99US-0135353.  
 PR 24-MAY-1999; 99US-0135629.  
 PR 25-MAY-1999; 99US-0136021.  
 PR 27-MAY-1999; 99US-0136392.  
 PR 28-MAY-1999; 99US-0136782.  
 PR 01-JUN-1999; 99US-0137222.  
 PR 03-JUN-1999; 99US-0137528.  
 PR 04-JUN-1999; 99US-0137528.  
 PR 07-JUN-1999; 99US-0137724.  
 PR 08-JUN-1999; 99US-0138094.  
 PR 10-JUN-1999; 99US-0138540.  
 PR 10-JUN-1999; 99US-0138847.



[illegible]

PR	12-AUG-1999;	9905-01483431;
PR	12-AUG-1999;	9905-0148565;
PR	13-AUG-1999;	9905-0148684;
PR	13-AUG-1999;	9905-0148368;
PR	16-AUG-1999;	9905-01494175;
PR	17-AUG-1999;	9905-01493175;
PR	18-AUG-1999;	9905-01484722;
PR	20-AUG-1999;	9905-0148722;
PR	20-AUG-1999;	9905-0149729;
PR	20-AUG-1999;	9905-0149902;
PR	23-AUG-1999;	9905-0149930;
PR	23-AUG-1999;	9905-0149930;
PR	25-AUG-1999;	9905-01505864;
PR	26-AUG-1999;	9905-0150884;
PR	27-AUG-1999;	9905-0151065;
PR	27-AUG-1999;	9905-0151086;
PR	27-AUG-1999;	9905-0151086;
PR	30-AUG-1999;	9905-01513038;
PR	31-AUG-1999;	9905-01514303;
PR	01-SEP-1999;	9905-01521369;
PR	07-SEP-1999;	9905-0152363;
PR	10-SEP-1999;	9905-0153070;
PR	13-SEP-1999;	9905-0153758;
PR	15-SEP-1999;	9905-01540188;
PR	16-SEP-1999;	9905-0154039;
PR	20-SEP-1999;	9905-01547479;
PR	22-SEP-1999;	9905-0155139;
PR	23-SEP-1999;	9905-01553869;
PR	24-SEP-1999;	9905-0155659;
PR	28-SEP-1999;	9905-01564588;
PR	29-SEP-1999;	9905-01565696;
PR	04-OCT-1999;	9905-01571177;
PR	05-OCT-1999;	9905-0157753;
PR	06-OCT-1999;	9905-0157865;
PR	07-OCT-1999;	9905-01580229;
PR	08-OCT-1999;	9905-01583232;
PR	12-OCT-1999;	9905-0158369;
PR	13-OCT-1999;	9905-0159293;
PR	13-OCT-1999;	9905-0159284;
PR	14-OCT-1999;	9905-0159295;
PR	14-OCT-1999;	9905-0159329;
PR	14-OCT-1999;	9905-0159330;
PR	14-OCT-1999;	9905-0159331;
PR	14-OCT-1999;	9905-0159637;
PR	14-OCT-1999;	9905-0159638;
PR	18-OCT-1999;	9905-0159584;
PR	21-OCT-1999;	9905-0160741;
PR	21-OCT-1999;	9905-0160761;
PR	21-OCT-1999;	9905-0160768;
PR	21-OCT-1999;	9905-0160770;
PR	21-OCT-1999;	9905-0160814;
PR	21-OCT-1999;	9905-0160815;
PR	21-OCT-1999;	9905-0160815;
PR	22-OCT-1999;	9905-0160981;
PR	22-OCT-1999;	9905-0160989;
PR	25-OCT-1999;	9905-0161404;
PR	25-OCT-1999;	9905-0161405;
PR	25-OCT-1999;	9905-0161406;
PR	26-OCT-1999;	9905-0161359;
PR	26-OCT-1999;	9905-0161360;
PR	26-OCT-1999;	9905-0161361;
PR	28-OCT-1999;	9905-0161920;
PR	28-OCT-1999;	9905-0161922;
PR	28-OCT-1999;	9905-0161993;
PR	29-OCT-1999;	9905-0162142;

Query Match	Similarity	25.8%	Score	259.8	DB	Length	774
Best Local	Similarity	78.2%	Pred No.	11e-62			
Matches	Conservative	0	Mismatches	87	Indels	0	Gaps
QY	59	atggagacatcacctgtgtgtgaaaagctctataccacaagaatgctcgaccctgaaccacaagag	118				
Db	1	atgggaagatccacacagctcgccgagaaagctcacatgaaacaagaagagctctggagcctaaga	60				

OY 119 gaagatcaaggcccatcacaataacgttcagtccaattgtaagctgcctggcttccc 178  
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||  
Db 61 gaagatcacgctcttglttgtatatccgaatcaaacgggttaaagttgcgcgatcccc 120  
  
OY 179 cccaagaagctgtcgctgcttagatgtgftaaaggtgtgcagattaagaatgataactac 238  
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||  
Db 121 ctctgcgcgcgtgttataccaaaaggtcggfcaaaggttgctgattgagatygatyaattat 160  
  
OY 239 ttgaagcctgtactttaagagaaggaanaattcaactgaagaaagaatgaagttatcatcaag 298  
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||  
Db 181 ctaagaccaagctctcaaaagagggagaattttacctgaagaagaaggtgacatcatcaag 240  
  
OY 239 ctccaagttactctggaacaacaatggtcatactgattcvtcygaagattaccagaagaaca 358  
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||  
Db 241 ctccaatagctgtgccggtgaacaacatggtcttctaataagctgyagattaccagaagaaca 300  
  
OY 359 gatatatgagataaagaactactctgtaaacacacacactcaaaagaagaagcttaacgaagga 418  
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||  
Db 301 gataacgagatccaagaactattgtaaacactcatctcaagaaggaagctctccaagcgttyg 360  
  
OY 419 attgattcaacaactatcgctcccttaatcaaacggcc 457  
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||  
Db 361 atygtatccaactctccaagctctgatalcaacgaatccgctc 399

PR	25-OCT-1999;	99US-0161405.
PR	25-OCT-1999;	99US-0161406.
PR	25-OCT-1999;	99US-0161389.
PR	26-OCT-1999;	99US-0161359.
PR	26-OCT-1999;	99US-0161360.
PR	26-OCT-1999;	99US-0161361.
PR	28-OCT-1999;	99US-0161920.
PR	28-OCT-1999;	99US-0161920.
PR	28-OCT-1999;	99US-0161993.
PR	29-OCT-1999;	99US-0162142.
<hr/>		
Query Match	25.2%;	Score 253.6; DB 21; Length 639;
Best Local Similarity	78.4%;	Pred. No. 5.2e-61;
Matches 304; Conservative	0;	Mismatches 84; Indels 0; Gaps
<hr/>		
OY	59 atggagacatcaccttgtttgaaaggctataccaacaaggctcgtgaccaagaag	118
Db	1 atgggaagatcaccatggttgtaaaggctccacatgacaagaagcatgtcaaga	60
OY	119 gaagataacgccctcaltaactatcatccgttcattcattgttgaaggctctgcgtccctc	178
Db	61 gaagatcaacgctcttatgtatatcatcgtatcatcattgtaaggctcttgcgtccctc	120
OY	179 cccaaagctgtgtggcgctgcttagatgtgtgtanagagtgtgcagattaagatgataaac	238
Db	121 cctaatacgtcttggtgtgtgtcgtctgtgtganaaagtgtagataaagatgatataac	160
OY	239 ttgagcgctgtactttaagagaggaattcactgaagaagaagatgagcttalcatcaag	298
Db	181 ctgcgtcctgtctttaaagcgtggaattcattcattgttgtgaaggsaatlcatgtc	240
OY	239 cttaacagttactttggaacaacaatgtgtcatgtattgtctgtgaagattaccagaaga	358
Db	241 cttaacagttatttttggcaaaaatgtgtcttgcattgtgtgtgaattaccggaagaac	300
OY	359 gataatgagataaagaactactgtgaacacacacatcaaagaagaagcttataagcaag	418
Db	301 gataatgagattaaaaattacttggaaacctcataataaagaagctcttaacgtggt	360
OY	419 attgattccacaactcattcgtccctca	446
Db	361 attgacccaaaactcaagtggtgatca	388
<hr/>		
RESULT 10		
ID AAD06456		
AAD06456 standard; cdna; 1087 BP.		
XX AAD06456;		
AC AAD06456;		
XX 10-AUG-2001 (first entry)		
DT XX Arabidopsis thaliana transcription factor G664 cDNA.		
DE XX Transcription factor; environmental stress tolerance; gene therapy;		
KW XX Plant structure; plant development; ss.		
RN XX Arabidopsis thaliana.		
OS XX Key Location/Qualifiers		
FT CDS 104..952		
FT FT /*tag= a		
XX PM /product= "Arabidopsis thaliana transcription factor"		
PM W0200136598-Al.		
PD 25-MAY-2001.		
XX 14-NOV-2000; 2000MO-US31458.		
XX PF 17-NOV-1999; 99US-0166228.		
PR PR 17-APR-2000; 2000US-0197889.		
PR PR 22-AUG-2000; 2000US-0227439.		
XX		



PR 18-JUN-1999; 99US-0139456.  
PR 18-JUN-1999; 99US-0139457.  
PR 18-JUN-1999; 99US-0139458.  
PR 18-JUN-1999; 99US-0139459.  
PR 18-JUN-1999; 99US-0139460.  
PR 18-JUN-1999; 99US-0139461.  
PR 18-JUN-1999; 99US-0139462.  
PR 18-JUN-1999; 99US-0139463.  
PR 18-JUN-1999; 99US-0139750.  
PR 18-JUN-1999; 99US-0139763.  
PR 21-JUN-1999; 99US-0139817.  
PR 22-JUN-1999; 99US-0139899.  
PR 23-JUN-1999; 99US-0140353.  
PR 23-JUN-1999; 99US-0140354.  
PR 24-JUN-1999; 99US-0140695.  
PR 28-JUN-1999; 99US-0140823.  
PR 29-JUN-1999; 99US-0140991.  
PR 30-JUN-1999; 99US-0141287.  
PR 01-JUL-1999; 99US-0141842.  
PR 01-JUL-1999; 99US-0142154.  
PR 02-JUL-1999; 99US-0142055.  
PR 06-JUL-1999; 99US-0142390.  
PR 08-JUL-1999; 99US-0142803.  
PR 09-JUL-1999; 99US-0142920.  
PR 12-JUL-1999; 99US-0142977.  
PR 13-JUL-1999; 99US-0143542.  
PR 14-JUL-1999; 99US-0143624.  
PR 15-JUL-1999; 99US-0144005.  
PR 16-JUL-1999; 99US-0144085.  
PR 16-JUL-1999; 99US-0144086.  
PR 19-JUL-1999; 99US-0144325.  
PR 19-JUL-1999; 99US-0144331.  
PR 19-JUL-1999; 99US-0144332.  
PR 19-JUL-1999; 99US-0144333.  
PR 19-JUL-1999; 99US-0144334.  
PR 19-JUL-1999; 99US-0144335.  
PR 20-JUL-1999; 99US-0144352.  
PR 20-JUL-1999; 99US-0144632.  
PR 21-JUL-1999; 99US-0144884.  
PR 21-JUL-1999; 99US-0144814.  
PR 21-JUL-1999; 99US-0145086.  
PR 21-JUL-1999; 99US-0145088.  
PR 22-JUL-1999; 99US-0145085.  
PR 22-JUL-1999; 99US-0145087.  
PR 22-JUL-1999; 99US-0145089.  
PR 22-JUL-1999; 99US-0145192.  
PR 23-JUL-1999; 99US-0145145.  
PR 23-JUL-1999; 99US-0145218.  
PR 23-JUL-1999; 99US-0145224.  
PR 26-JUL-1999; 99US-0145276.  
PR 27-JUL-1999; 99US-0145913.  
PR 27-JUL-1999; 99US-0145918.  
PR 27-JUL-1999; 99US-0145919.  
PR 28-JUL-1999; 99US-0145951.  
PR 02-AUG-1999; 99US-0146386.  
PR 02-AUG-1999; 99US-0146388.  
PR 02-AUG-1999; 99US-0146389.  
PR 03-AUG-1999; 99US-0147038.  
PR 04-AUG-1999; 99US-0147204.  
PR 04-AUG-1999; 99US-0147302.  
PR 05-AUG-1999; 99US-0147192.  
PR 05-AUG-1999; 99US-0147260.  
PR 06-AUG-1999; 99US-0147303.  
PR 06-AUG-1999; 99US-0147416.  
PR 09-AUG-1999; 99US-0147493.  
PR 09-AUG-1999; 99US-0147935.  
PR 10-AUG-1999; 99US-0148171.  
PR 11-AUG-1999; 99US-0148319.  
PR 12-AUG-1999; 99US-0148341.  
PR 13-AUG-1999; 99US-0148565.  
PR 13-AUG-1999; 99US-0148684.  
PR 16-AUG-1999; 99US-0149368.  
PR 17-AUG-1999; 99US-0149175.

PR 18-AUG-1999; 99US-0149426.  
PR 20-AUG-1999; 99US-0149722.  
PR 20-AUG-1999; 99US-0149723.  
PR 20-AUG-1999; 99US-0149929.  
PR 23-AUG-1999; 99US-0149902.  
PR 23-AUG-1999; 99US-0149930.  
PR 23-AUG-1999; 99US-0150566.  
PR 26-AUG-1999; 99US-0150884.  
PR 27-AUG-1999; 99US-0151065.  
PR 27-AUG-1999; 99US-0151066.  
PR 27-AUG-1999; 99US-0151080.  
PR 30-AUG-1999; 99US-0151303.  
PR 31-AUG-1999; 99US-0151438.  
PR 01-SEP-1999; 99US-0151930.  
PR 07-SEP-1999; 99US-0152363.  
PR 10-SEP-1999; 99US-0153070.  
PR 13-SEP-1999; 99US-0153758.  
PR 15-SEP-1999; 99US-0154018.  
PR 16-SEP-1999; 99US-0154039.  
PR 20-SEP-1999; 99US-0154779.  
PR 22-SEP-1999; 99US-0155139.  
PR 23-SEP-1999; 99US-0155486.  
PR 24-SEP-1999; 99US-0155659.  
PR 28-SEP-1999; 99US-0156458.  
PR 29-SEP-1999; 99US-0156596.  
PR 04-OCT-1999; 99US-0157717.  
PR 05-OCT-1999; 99US-0157753.  
PR 06-OCT-1999; 99US-0157865.  
PR 07-OCT-1999; 99US-0158029.  
PR 08-OCT-1999; 99US-0158232.  
PR 12-OCT-1999; 99US-0158369.  
PR 13-OCT-1999; 99US-0159293.  
PR 13-OCT-1999; 99US-0159294.  
PR 13-OCT-1999; 99US-0159295.  
PR 14-OCT-1999; 99US-0159329.  
PR 14-OCT-1999; 99US-0159330.  
PR 14-OCT-1999; 99US-0159331.  
PR 14-OCT-1999; 99US-0159637.  
PR 14-OCT-1999; 99US-0159638.  
PR 18-OCT-1999; 99US-0159584.  
PR 21-OCT-1999; 99US-0160741.  
PR 21-OCT-1999; 99US-0160767.  
PR 21-OCT-1999; 99US-0160768.  
PR 21-OCT-1999; 99US-0160770.  
PR 21-OCT-1999; 99US-0160814.  
PR 21-OCT-1999; 99US-0160815.  
PR 22-OCT-1999; 99US-0160980.  
PR 22-OCT-1999; 99US-0160981.  
PR 22-OCT-1999; 99US-0160989.  
PR 25-OCT-1999; 99US-0161404.  
PR 25-OCT-1999; 99US-0161405.  
PR 25-OCT-1999; 99US-0161406.  
PR 26-OCT-1999; 99US-0161359.  
PR 26-OCT-1999; 99US-0161360.  
PR 26-OCT-1999; 99US-0161361.  
PR 26-OCT-1999; 99US-0161920.  
PR 28-OCT-1999; 99US-0161992.  
PR 28-OCT-1999; 99US-0161993.  
PR 29-OCT-1999; 99US-0162142.

Query Match 24.6%; Score 247.8; DB 21; Length 1062;  
Best Local Similarity 76.7%; Pred. No. 2.7e-59;  
Matches 303; Conservative 0; Mismatches 92; Indels 0; Gaps 0;

QY 59 atgggaagcattactctgttggaaaggctataccacaagaagtgccctggaccaaagag 118  
DB 214 atgggaaggtaccctgtgtggaagaagctacacaaagaagacatggacgaagaagaa 273  
QY 119 gaagatcaacgcctcatcaactatccgtgtccatgtgaaagctgtgcgttccctc 178  
DB 274 gaggaagagagagctgtgccttaagctcatgtgaagagagctgtgcgtgagatctcc 333  
QY 179 cccaaagctgtggtcttaagatgtgttaagagtgcaagattagatgataactac 238

```

Db      334 cccaagcgcgcgactctctcgtgcaagagctgcgctccgttgatcaactat 393
Oy      239 ttgagcctgactatgaagagaattcactggaagaagaatgattatcaag 298
Db      394 ctccgctgactctgaagctggaacttcacccgaggaagaagaccatcatcaag 453
Oy      299 ctccagcttactctggaacaatgctgattgattctggaagattaccaggaagaca 358
Db      454 ctccatagctctctggaacaatgctgattgattgctcgaggatattaccgggaagaca 513
Oy      359 gataatgataaagaactactggaacacacacatcaaaagaagcttataagcagaga 418
Db      514 gataacgagataaagaactattggaacacgcatatacgaagaagcttataaacaagagg 573
Oy      419 attgaccacaactatgctctctcaatcaac 453
Db      574 attgaccacaagctatgaccacatccaagaatc 608

```

## RESULT 12

AAC56996  
ID AAC56996 standard; DNA; 631 BP.

AC AAC56996;

DT 25-JAN-2001 (first entry)

XX Pinus radiata transcription factor DNA sequence #442.

XX Plant; transcription factor; gene expression; eucalyptus; pine; acacia;  
XX poplar; sweetgum; teak; mahogany; bZIP; G-box binding factor;  
XX basic helix-loop-helix zipper; homeotic; homeodomain; homeobox; MADS;  
XX homeodomain zipper; LIM domain; AP2; ERBS; zinc finger domain;  
XX type 2 Cys2His2; CCAAT box element; MYB; ss.

OS Pinus radiata.

PN WO200053724-A2.

PD 14-SEP-2000.

PE 09-MAR-2000; 2000MO-US06112.

PR 11-MAR-1999; 99US-0266513.

PR 18-AUG-1999; 99US-0149485.

XX (GENE-) GENESIS RES & DEV CORP LTD.

PA (FLET-) FLETCHER CHALLENGE FORESTS LTD.

XX Wood M, McGrath A, Shenk MA, Glenn M;

PI WPI; 2000-579369/54.

XX New isolated polynucleotide encoding a plant transcription factor for  
XX producing a plant e.g. a woody plant, preferably eucalyptus or pine,  
XX having modified gene expression or modified activity of a polypeptide

XX Claim 1; Page 562; 747pp; English.

XX The present invention relates to novel plant transcription factors from  
XX Eucalyptus grandis or Pinus radiata. The present sequence is the coding  
XX sequence for one such transcription factor. The transcription factor may  
XX be used to produce a plant having modified gene expression such as a  
XX woody plant e.g. a eucalyptus, pine, acacia, poplar, sweetgum, teak, or  
XX mahogany species or to modify the activity of a polypeptide in a plant.  
XX The transcription factors of the present invention are members from the  
XX following families of regulatory proteins: bZIP, bZIP family of G-box  
XX binding factors, basic helix-loop-helix zipper, LIM domain, AP2  
XX homeotic/homeodomain/homeobox/MADS, homeodomain zipper, LIM domain, AP2  
XX and ERBS, zinc finger domains of type 2 Cys2His2, CCAAT box elements  
XX and MYB.

XX  
SQ Sequence 631 BP; 165 A; 185 C; 153 G; 128 T; 0 other;

Query Match 24.4%; Score 245.2; DB 21; Length 631;  
Best Local Similarity 73.5%; Pred. No. 1.1e-58;  
Matches 313; Conservative 0; Mismatches 113; Indels 0; Gaps 0;

```

Oy      43 ttgattttgtaactgattggaacgatacctgtttggaagaagctatcaacaagaag 102
Db      91 tctgtgtatctgaagatgaggagatccctcctgttggaaaagctcataaacaagaag 150
Oy      103 tgcctggaaccaagaagagaatcaacgctcatcaatcacctcgttccatggtgaag 162
Db      151 ggcgtggaaccaagaagagaagagatcgatccgcatcgcccatctgaaactcagcgcaag 210
Oy      163 ctgctggcgttccctcccaagaagctgctggtgcttagatgtgtgaagttagat 222
Db      211 ttgctggcgtcgtctcccaagaagcgcgaaggtctgattgctgcyggaaggttgaagct 270
Oy      223 aagatgataaactactgagcctgatactttaagagaagaattcactggaagaaga 282
Db      271 ccgattgataaactactcgtctctatctcgaagcgtggaactctcagaagaaga 330
Oy      283 tgagcttatacataagcttcacagttacttggaaacaatgltcatgtattgttgaag 342
Db      331 cgaactcgtcatcaaaactccatccatccatcgtgcaacaagtggtcttattgagcag 390
Oy      343 attacagaagaagaacagataatgagataaagaactacttggaaacacacatcaagaaga 402
Db      391 attgcccgggcgagcagacacagataaagaactacttggaaactcatcaagaaga 450
Oy      403 gctttaaagcagagattgattgacacaaactcgtctctcaatcaaaagccaatc 462
Db      451 attgctaaacagaggactcgaccctccagctccatgcccctcggaagccgcaacaag 510
Oy      463 caacac 468
Db      511 caacac 516

```

## RESULT 13

AAC57348  
ID AAC57348 standard; DNA; 1150 BP.

AC AAC57348;

DT 25-JAN-2001 (first entry)

XX Pinus radiata transcription factor DNA sequence #703.

XX Plant; transcription factor; gene expression; eucalyptus; pine; acacia;  
XX poplar; sweetgum; teak; mahogany; bZIP; G-box binding factor;  
XX basic helix-loop-helix zipper; homeotic; homeodomain; homeobox; MADS;  
XX homeodomain zipper; LIM domain; AP2; ERBS; zinc finger domain;  
XX type 2 Cys2His2; CCAAT box element; MYB; ss.

OS Pinus radiata.

PN WO200053724-A2.

PD 14-SEP-2000.

PE 09-MAR-2000; 2000MO-US06112.

PR 11-MAR-1999; 99US-0266513.

PR 18-AUG-1999; 99US-0149485.

XX (GENE-) GENESIS RES & DEV CORP LTD.

PA (FLET-) FLETCHER CHALLENGE FORESTS LTD.

XX Wood M, McGrath A, Shenk MA, Glenn M;

PI WPI; 2000-579369/54.

DR WPI: 2000-579369/54.

XX New isolated polynucleotide encoding a plant transcription factor for  
PT producing a plant e.g. a woody plant, preferably eucalyptus or pine,  
PT having modified gene expression or modified activity of a polypeptide  
PT

PS Claim 1; Page 653; 747pp; English.

XX The present invention relates to novel plant transcription factors from  
CC Eucalyptus grandis or Pinus radiata. The present sequence is the coding  
CC sequence for one such transcription factor. The transcription factor may  
CC be used to produce a plant having modified gene expression such as a  
CC woody plant e.g. a eucalyptus, pine, acacia, poplar, sweetgum, teak, or  
CC mahogany species or to modify the activity of a polypeptide in a plant.  
CC The transcription factors of the present invention are members from the  
CC following families of regulatory proteins: bZIP, bZIP family of G-box  
CC binding factors, basic helix-loop-helix zipper,  
CC homeotic/homeodomain/homeobox/MADS, homeodomain zipper, LIM domain, AP2  
CC and EREBs, zinc finger domains of type 2 Cys2His2, CCAAT box elements  
CC and MYB.

XX Sequence 1150 BP; 320 A; 287 C; 277 G; 266 T; 0 other;

Query Match 24.4%; Score 245.2; DB 21; Length 1150;

Best Local Similarity 73.5%; Pred. No. 1.5e-58;  
Matches 313; Conservative 0; Mismatches 113; Indels 0; Gaps 0;

QY 43 ttgatttgaactgaatgagcagatcactgtgttgaaagctatacaacaaga 102  
DB 47 tctgtgtatgtgaagatgagagagatcctcctgtgtgaaagctatacaacaaga 106

QY 103 tgcctggaccacaaaggagagatcaacgctcactcaactacatccgtgtccatgtgaaag 162  
DB 107 ggcgtggaccacaaaggagagatcagatcgtcctcactcagccacattcgaactcagcgcaag 166

QY 163 ctgcgcggtccctcccccacaaagctgctgagctagatgtgtgaaagctcagatt 222  
DB 167 ttgcgcggtccctcccccacaaagctgagctgagctgctgagaaagctcagatt 226

QY 223 aagatgataaactactgtgagcctgtactcttaagaagaagaatttcactcagaagaaga 282  
DB 227 ccgattgataaactactgtgagcctgtactctgtgaagcgtggaactctcagaagaagaaga 286

QY 283 tgaagctatcaacagctcagcttactgtgaaacaagaatgtgcatgtgtgagag 342  
DB 287 cgaactcgtcactcaaacactcactcctcactcgcgaacaagtgtctcttatttgagagag 346

QY 343 attaccagaaagaaagcagaatgagataaagaactactgtgaaacacacacacacacacaa 402  
DB 347 attgcgcgggagcagagcagaacagagataaagaactactgtgaaatcactcacaagaaga 406

QY 403 gcttaaaagcagaagaattgtatcacaacaactcgtcctctcactcaatcaaacggccaat 462  
DB 407 attgctaaacagggagctgagcccccagtcctcagccctcgcgcagcgcaacaacag 466

QY 463 caacac 468  
DB 467 caacac 472

RESULT 14

AAD06466  
ID AAD06466 standard; cDNA; 918 BP;

XX AAD06466;

XX 10-AUG-2001 (first entry)

XX Arabidopsis thaliana transcription factor G664 homologue, G255 cDNA.

XX Transcription factor; environmental stress tolerance; gene therapy;

KW plant structure; plant development; ss.

OS Arabidopsis thaliana.

XX Key Location/Qualifiers

FT CDS 30..839

FT /tag= a

FT /product= "Arabidopsis thaliana transcription factor

FT homologue"

XX WO200136598-A1.

XX 25-MAY-2001.

XX 14-NOV-2000; 2000WO-US31458.

XX 17-NOV-1999; 99US-0166228.

XX 17-APR-2000; 2000US-0197899.

XX 22-AUG-2000; 2000US-0227439.

XX (MEND-) MENDEL BIOTECHNOLOGY INC.

XX (PINE/) PINEDA O.

XX (YUG/) YU G.

XX (CREE/) CREELMAN R.

XX (RICH/) RIECHMANN J L.

XX (HEAR/) HEARD J.

XX (RATC/) RATCLIFFE O.

XX (REUB/) REUBER L.

XX (KEDD/) KEDDIE J.

XX pineda O, Yu G, Creelman R, Riechmann JL, Heard J, Ratcliffe O,

XX Reuber L, Keddie J;

XX WPI: 2001-336000/35.

XX P-PSDB: AAE02486.

XX Nucleic acids encoding plant transcription factor polypeptides, useful

XX for altering the environmental stress tolerance characteristics of

XX plants -

XX Claim 4; Page 102-103; 116pp; English.

XX The present sequence is a cDNA encoding Arabidopsis thaliana

XX transcription factor homologue. This novel transcription factor

XX is useful for modifying a plant's phenotype in desirable ways, such as

XX modifying a plant's environmental stress. The transcription factor is

XX encoded by environmental stress tolerance gene derived from

XX Arabidopsis thaliana. The transcription factors and the genes encoding

XX them are used to alter the structure and developmental characteristics

XX of plants such as soybean, wheat, corn, potato, cotton, rice, oilseed

XX rape, sunflower, alfalfa, sugarcane, turf, banana, blackberry, papaya,

XX strawberry, raspberry, cantaloupe, carrot, cauliflower, coffee, onion,

XX cucumber, eggplant, grapes, honey dew, lettuce, mango, melon, blueberry,

XX peas, peppers, pineapple, spinach, squash, sweet corn, tobacco, tomato,

XX watermelon, roseaceous fruits and/or vegetable brassicas. These sequences

XX are also used for modifying traits associated with environmental stress

XX tolerance, such as freezing, chilling, heat, drought, water saturation,

XX salt, photoconditions, radiation and ozone. The transcription factors

XX are used in gene therapy.

XX Sequence 918 BP; 307 A; 155 C; 217 G; 239 T; 0 other;

Query Match 23.4%; Score 235.2; DB 22; Length 918;

Best Local Similarity 73.5%; Pred. No. 8.4e-56;  
Matches 300; Conservative 0; Mismatches 108; Indels 0; Gaps 0;

QY 40 ttgatttgaactgaatgagcagatcactgtgttgaaagctatacaacaaga 99  
DB 11 tcatcagaagaagagagatcatgagagatcctcctgtgctgcggaagaacatgacaa 70

QY 100 aggtgcttgacacaaagagagatcaacgctcactcaactacatccgtgtcactgag 159  
DB 11 tcatcagaagaagagagatcatgagagatcctcctgtgctgcggaagaacatgacaa 70



Db 71 aggtgcttgagactaagaagaagatgagactagctcttcatcaatcaagtcacggtga 130  
Qy 160 aggtctgctgagcttcccccacaaagctgctgctgcttagatgtgtgaaggttcag 219  
Db 131 aggttgttgagcatctctccacagccgctgtctctccttgctggtgtaaaagcgtccg 190  
Qy 220 attaagatgtaataactctctgaagccttgatcttaagagaagaatttcacaaaga 279  
Db 191 tcttcggaggatcaatctccgcacgtcatctcaaaagaggaacttcaacatgatga 250  
Qy 280 agatgagctttatcaacagcttcaacagttacttgaacaacatgttcattgtctg 339  
Db 251 agatgaacttcatcaacagcttcatagcctccttagcaacaagtgtcttgatgcg 310  
Qy 340 aagattaccaggaagaacagatgagataaagaacttctggaacacacacacacaaag 399  
Db 311 gagattaccctggaagaacagatgagatcaagaactctggaacacacacacataaag 370  
Qy 400 aaagcttaagcagaggaattgatccacaacatcgtcctcctca 447  
Db 371 gaagctttgagcaagagattgatccacagcactcatagaggatca 418  
RESULT: 15  
AAC2479  
ID AAC2479 standard; DNA; 839 BP.  
XX AAC2479;  
AC AAC2479;  
XX  
DT 17-OCT-2000 (first entry)  
XX  
DE Arabidopsis thaliana DNA fragment SEQ ID NO: 35715.  
XX  
KW Hybridisation assay; genetic mapping; gene expression control;  
KW protein identification; signal transduction pathway;  
KW metabolic pathway; promoter; termination sequence; ss.  
XX  
OS Arabidopsis thaliana.  
XX  
PN EP1033405-A2.  
XX  
PD 06-SEP-2000.  
XX  
PF 25-FEB-2000; 2000EP-0301439.  
XX  
PR 25-FEB-1999; 99US-0121825.  
PR 05-MAR-1999; 99US-0123180.  
PR 09-MAR-1999; 99US-0123548.  
PR 23-MAR-1999; 99US-0125788.  
PR 25-MAR-1999; 99US-0126264.  
PR 29-MAR-1999; 99US-0126785.  
PR 01-APR-1999; 99US-0127462.  
PR 06-APR-1999; 99US-0128234.  
PR 08-APR-1999; 99US-0128714.  
PR 16-APR-1999; 99US-0129845.  
PR 19-APR-1999; 99US-0130077.  
PR 21-APR-1999; 99US-0130449.  
PR 23-APR-1999; 99US-0130510.  
PR 23-APR-1999; 99US-0130891.  
PR 28-APR-1999; 99US-0131449.  
PR 30-APR-1999; 99US-0132048.  
PR 30-APR-1999; 99US-0132407.  
PR 04-MAY-1999; 99US-0132484.  
PR 05-MAY-1999; 99US-0132485.  
PR 06-MAY-1999; 99US-0132486.  
PR 06-MAY-1999; 99US-0132487.  
PR 07-MAY-1999; 99US-0132863.  
PR 11-MAY-1999; 99US-0134256.  
PR 14-MAY-1999; 99US-0134218.  
PR 14-MAY-1999; 99US-0134219.  
PR 14-MAY-1999; 99US-0134221.  
PR 14-MAY-1999; 99US-0134370.  
PR 18-MAY-1999; 99US-0134768.  
PR 18-MAY-1999; 99US-0134768.

PR 19-MAY-1999; 99US-0134941.  
PR 20-MAY-1999; 99US-0135124.  
PR 21-MAY-1999; 99US-0135353.  
PR 24-MAY-1999; 99US-0135629.  
PR 25-MAY-1999; 99US-0136021.  
PR 27-MAY-1999; 99US-0136392.  
PR 28-MAY-1999; 99US-0136782.  
PR 01-JUN-1999; 99US-0137222.  
PR 03-JUN-1999; 99US-0137528.  
PR 04-JUN-1999; 99US-0137502.  
PR 07-JUN-1999; 99US-0137724.  
PR 08-JUN-1999; 99US-0138094.  
PR 10-JUN-1999; 99US-0138540.  
PR 14-JUN-1999; 99US-0138817.  
PR 16-JUN-1999; 99US-0139452.  
PR 17-JUN-1999; 99US-0139453.  
PR 18-JUN-1999; 99US-0139454.  
PR 18-JUN-1999; 99US-0139455.  
PR 18-JUN-1999; 99US-0139456.  
PR 18-JUN-1999; 99US-0139457.  
PR 18-JUN-1999; 99US-0139458.  
PR 18-JUN-1999; 99US-0139459.  
PR 18-JUN-1999; 99US-0139460.  
PR 18-JUN-1999; 99US-0139461.  
PR 18-JUN-1999; 99US-0139462.  
PR 18-JUN-1999; 99US-0139463.  
PR 18-JUN-1999; 99US-0139750.  
PR 18-JUN-1999; 99US-0139763.  
PR 21-JUN-1999; 99US-0139817.  
PR 22-JUN-1999; 99US-0139899.  
PR 23-JUN-1999; 99US-0140353.  
PR 23-JUN-1999; 99US-0140354.  
PR 24-JUN-1999; 99US-0140695.  
PR 28-JUN-1999; 99US-0140823.  
PR 29-JUN-1999; 99US-0140991.  
PR 30-JUN-1999; 99US-0141287.  
PR 01-JUL-1999; 99US-0141842.  
PR 01-JUL-1999; 99US-0142154.  
PR 02-JUL-1999; 99US-0142055.  
PR 06-JUL-1999; 99US-0142390.  
PR 09-JUL-1999; 99US-0142803.  
PR 09-JUL-1999; 99US-0142920.  
PR 12-JUL-1999; 99US-0142977.  
PR 13-JUL-1999; 99US-0143542.  
PR 14-JUL-1999; 99US-0143624.  
PR 15-JUL-1999; 99US-0144005.  
PR 16-JUL-1999; 99US-0144085.  
PR 16-JUL-1999; 99US-0144086.  
PR 19-JUL-1999; 99US-0144325.  
PR 19-JUL-1999; 99US-0144331.  
PR 19-JUL-1999; 99US-0144332.  
PR 19-JUL-1999; 99US-0144333.  
PR 19-JUL-1999; 99US-0144334.  
PR 19-JUL-1999; 99US-0144335.  
PR 20-JUL-1999; 99US-0144352.  
PR 20-JUL-1999; 99US-0144632.  
PR 20-JUL-1999; 99US-0144884.  
PR 21-JUL-1999; 99US-0144814.  
PR 21-JUL-1999; 99US-0145086.  
PR 21-JUL-1999; 99US-0145088.  
PR 22-JUL-1999; 99US-0145085.  
PR 22-JUL-1999; 99US-0145087.  
PR 22-JUL-1999; 99US-0145089.  
PR 22-JUL-1999; 99US-0145192.  
PR 23-JUL-1999; 99US-0145145.  
PR 23-JUL-1999; 99US-0145218.  
PR 23-JUL-1999; 99US-0145224.  
PR 26-JUL-1999; 99US-0145276.  
PR 27-JUL-1999; 99US-0145913.  
PR 27-JUL-1999; 99US-0145918.  
PR 27-JUL-1999; 99US-0145919.

PR 28-JUL-1999; 99US-014551.  
PR 02-AUG-1999; 99US-0146386.  
PR 02-AUG-1999; 99US-0146388.  
PR 02-AUG-1999; 99US-0146389.  
PR 03-AUG-1999; 99US-0147038.  
PR 04-AUG-1999; 99US-0147204.  
PR 04-AUG-1999; 99US-0147302.  
PR 05-AUG-1999; 99US-0147302.  
PR 05-AUG-1999; 99US-0147360.  
PR 06-AUG-1999; 99US-0147303.  
PR 06-AUG-1999; 99US-0147416.  
PR 09-AUG-1999; 99US-0147493.  
PR 09-AUG-1999; 99US-0147935.  
PR 10-AUG-1999; 99US-0148171.  
PR 11-AUG-1999; 99US-0148319.  
PR 12-AUG-1999; 99US-0148341.  
PR 13-AUG-1999; 99US-0148365.  
PR 13-AUG-1999; 99US-0148684.  
PR 16-AUG-1999; 99US-0149368.  
PR 17-AUG-1999; 99US-0149175.  
PR 18-AUG-1999; 99US-0149426.  
PR 20-AUG-1999; 99US-0149722.  
PR 20-AUG-1999; 99US-0149723.  
PR 20-AUG-1999; 99US-0149929.  
PR 23-AUG-1999; 99US-0149902.  
PR 23-AUG-1999; 99US-0149930.  
PR 25-AUG-1999; 99US-0150566.  
PR 26-AUG-1999; 99US-0150884.  
PR 27-AUG-1999; 99US-0151065.  
PR 27-AUG-1999; 99US-0151066.  
PR 27-AUG-1999; 99US-0151080.  
PR 30-AUG-1999; 99US-0151303.  
PR 31-AUG-1999; 99US-0151438.  
PR 01-SEP-1999; 99US-0151930.  
PR 07-SEP-1999; 99US-0152363.  
PR 10-SEP-1999; 99US-0153070.  
PR 13-SEP-1999; 99US-0153758.  
PR 15-SEP-1999; 99US-0154018.  
PR 16-SEP-1999; 99US-0154039.  
PR 20-SEP-1999; 99US-0154779.  
PR 22-SEP-1999; 99US-0155139.  
PR 23-SEP-1999; 99US-0155486.  
PR 24-SEP-1999; 99US-0155659.  
PR 28-SEP-1999; 99US-0156458.  
PR 29-SEP-1999; 99US-0156596.  
PR 04-OCT-1999; 99US-0157117.  
PR 05-OCT-1999; 99US-0157753.  
PR 06-OCT-1999; 99US-0157865.  
PR 07-OCT-1999; 99US-0158029.  
PR 08-OCT-1999; 99US-0158232.  
PR 12-OCT-1999; 99US-0158369.  
PR 13-OCT-1999; 99US-0159293.  
PR 13-OCT-1999; 99US-0159294.  
PR 13-OCT-1999; 99US-0159295.  
PR 14-OCT-1999; 99US-0159329.  
PR 14-OCT-1999; 99US-0159330.  
PR 14-OCT-1999; 99US-0159331.  
PR 14-OCT-1999; 99US-0159637.  
PR 14-OCT-1999; 99US-0159638.  
PR 18-OCT-1999; 99US-0159584.  
PR 21-OCT-1999; 99US-0160741.  
PR 21-OCT-1999; 99US-0160767.  
PR 21-OCT-1999; 99US-0160768.  
PR 21-OCT-1999; 99US-0160770.  
PR 21-OCT-1999; 99US-0160814.  
PR 21-OCT-1999; 99US-0160815.  
PR 22-OCT-1999; 99US-0160980.  
PR 22-OCT-1999; 99US-0160981.  
PR 22-OCT-1999; 99US-0160989.  
PR 25-OCT-1999; 99US-0161404.  
PR 25-OCT-1999; 99US-0161405.  
PR 25-OCT-1999; 99US-0161406.  
PR 26-OCT-1999; 99US-0161359.

PR 26-OCT-1999; 99US-0161360.  
PR 26-OCT-1999; 99US-0161361.  
PR 28-OCT-1999; 99US-0161920.  
PR 28-OCT-1999; 99US-0161992.  
PR 28-OCT-1999; 99US-0161993.  
PR 29-OCT-1999; 99US-0162142.

Query Match 23.2%; Score 233.8; DB 21; Length 838;  
Best Local Similarity 75.1%; Pred. No. 2e-55;  
Matches 292; Conservative 0; Mismatches 97; Indels 0; Gaps 0;

QY 59 atggagcagatccttctgtgaaaggctacacaaaggctcctggacaaagg  
Db 30 atggagcagatccttctgtgcgagaagacacagaaaggctcctggacaaagg  
QY 119 gaaatcaagccctcaactacatccgtgcagatggaagcctgcggctccctc  
Db 90 gaaatgagagactagctcttacaacaaagctccagtgaaagctgttcgagatcct  
QY 179 cccaagctgtggtccttagatgtgttaagattgcaagttaaagatggaactac  
Db 150 cctagagcctgtgtctcctcctcctgcgtgaaagctgcgtcctcgtggaactac  
QY 239 ttgagcctgtactcttaagagagaaattcactgaaagaaagatgagttacatcaag  
Db 210 ctcgacctgtactcacaagagaaagaaacttacaatgaaagatgaaactatcatcaag  
QY 299 cttaacagttactcttgaaacaatggtcatatgattgtctgaaagattacaaagaaaca  
Db 270 cttaacagttactcttgaaacaatggtcatatgattgtctgaaagattacaaagaaaca  
QY 359 gataatgagataaagaaactacttgaaacacacacacataaagaagaagctttaagcagaaga  
Db 330 gataacgagataaagaaactacttgaaacacacacacataaagaagaagctttaagcagaaga  
QY 419 atgattcacaactcactgcctctcaaa 447  
Db 390 atgattcagcactcactagagggatcaaa 418

Search completed: January 23, 2002, 20:25:21  
Job time: 3469 sec

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 23, 2002, 18:48:37 ; Search time 90.74 Seconds

(without alignments)  
2510.876 Million cell updates/sec

Title: US-09-453-387A-1

Perfect score: 1006  
Sequence: 1 taaccacgtattcttc.....aatgctaattataaaaa 1006

Scoring table: IDENTITY\_NUC  
Gapop 10.0, Gapext 1.0

Searched: 351203 seqs, 11323899 residues

Total number of hits satisfying chosen parameters: 702406

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database : Issued\_Patents\_NA.\*

1: /cgn2\_6/ptodata/2/ina/5A.COMB.seq.\*  
2: /cgn2\_6/ptodata/2/ina/5B.COMB.seq.\*  
3: /cgn2\_6/ptodata/2/ina/6A.COMB.seq.\*  
4: /cgn2\_6/ptodata/2/ina/6B.COMB.seq.\*  
5: /cgn2\_6/ptodata/2/ina/PCrus.COMB.seq.\*  
6: /cgn2\_6/ptodata/2/ina/Backfiles1.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	162.2	16.1	1344	US-08-722-626B-1	Sequence 1, Appl
2	109.6	10.9	2220	US-08-997-251-1	Sequence 1, Appl
3	109.2	10.9	2352	US-08-997-251-3	Sequence 3, Appl
4	79.2	7.9	1035	PCT-US93-06251-90	Sequence 90, Appl
5	79.2	7.9	3225	US-08-306-691B-45	Sequence 45, Appl
6	79.2	7.9	3225	PCT-US93-06251-91	Sequence 91, Appl
7	79.2	7.9	3230	PCT-US93-06251-89	Sequence 89, Appl
8	69.8	6.9	785	US-09-008-979A-6	Sequence 6, Appl
9	69.8	6.9	785	US-09-460-618-6	Sequence 6, Appl
10	69.8	6.9	2638	US-08-306-691B-46	Sequence 46, Appl
11	63.6	6.3	469	US-09-008-979B-2	Sequence 2, Appl
12	63.6	6.3	469	US-09-460-618-2	Sequence 2, Appl
13	56	5.6	4059	US-08-485-139-1	Sequence 1, Appl
14	56	5.6	4059	US-08-750-357-1	Sequence 1, Appl
15	56	5.6	4824	US-08-485-139-5	Sequence 5, Appl
16	56	5.6	4824	US-08-750-357-5	Sequence 5, Appl
17	49.6	4.9	7218	US-08-232-463-14	Sequence 14, Appl
18	43.2	4.3	282	US-08-814-030-2	Sequence 2, Appl
19	35	3.5	346	US-09-040-984-3	Sequence 3, Appl
20	33.6	3.3	1750	US-08-415-823-1	Sequence 1, Appl
21	33.6	3.3	1750	US-09-086-662-1	Sequence 1, Appl
22	33.4	3.3	2743	US-08-317-707-1	Sequence 1, Appl
23	33	3.3	289	US-09-007-005-17	Sequence 17, Appl
24	33	3.3	289	US-09-244-796-17	Sequence 17, Appl
25	33	3.3	2749	US-08-344-695-1	Sequence 1, Appl
26	31.6	3.1	256	US-08-905-223-66	Sequence 66, Appl
27	31.6	3.1	1215	US-08-642-274D-35	Sequence 35, Appl

28	31.6	3.1	1215	US-08-952-014C-35	Sequence 35, Appl
29	31.6	3.1	7587	US-08-378-313-22	Sequence 22, Appl
30	31.6	3.1	48974	US-08-920-422-17	Sequence 17, Appl
31	31.4	3.1	2247	US-08-760-615-5	Sequence 5, Appl
32	31.2	3.1	2290	5312912-1	Patent No. 5312912
33	31	3.1	2317	US-08-749-522-5	Sequence 5, Appl
34	31	3.1	3574	US-08-467-504-3	Sequence 3, Appl
35	30.8	3.1	344	US-08-623-906A-8	Sequence 8, Appl
36	30.8	3.1	864	US-08-906-769-138	Sequence 138, App
37	30.8	3.1	864	US-08-906-616-138	Sequence 138, App
38	30.8	3.1	864	US-08-639-075A-138	Sequence 138, App
39	30.8	3.1	864	US-09-012-431-138	Sequence 138, App
40	30.8	3.1	864	US-09-012-692-138	Sequence 138, App
41	30.8	3.1	864	US-08-906-613-138	Sequence 138, App
42	30.8	3.1	2435	US-09-306-593-1	Sequence 1, Appl
43	30.6	3.0	2160	US-08-902-853-4	Sequence 4, Appl
44	30.6	3.0	2383	US-08-248-466B-13	Sequence 13, Appl
45	30.6	3.0	5336	US-09-102-528-11	Sequence 11, Appl

## ALIGNMENTS

RESULT 1  
US-08-722-626B-1  
Sequence 1, Application US/08722626B  
Patent No. 5939601  
GENERAL INFORMATION:  
APPLICANT: Yang, Yimong  
APPLICANT: Kiessig, Daniel, F.  
TITLE OF INVENTION: NOVEL GENES ASSOCIATED WITH ENHANCED  
TITLE OF INVENTION: NOVEL GENES ASSOCIATED WITH ENHANCED  
NUMBER OF SEQUENCES: 9  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Dann, Dorfman, Herrell and Skillman  
STREET: 1601 Market Street Suite 720  
CITY: Philadelphia  
STATE: PA  
COUNTRY: USA  
ZIP: 19103-2307  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FastSeq Version 1.5  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/722, 626B  
FILING DATE: 27-SEP-1996  
CLASSIFICATION: 800  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER:  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Pat Hagan  
REGISTRATION NUMBER: 27,643  
REFERENCE/DOCKET NUMBER: 97-0010  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 215 563-4100  
TELEFAX: 215 563-4044  
TELEX:  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1344 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: CDNA  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
FRAGMENT TYPE:  
ORIGINAL SOURCE:  
FEATURE:  
NAME/KEY: Coding Sequence

LOCATION: 148...981  
OTHER INFORMATION:  
US-08-722-626B-1

Query Match 16.1%; Score 162.2; DB 2; Length 1344;  
Best Local Similarity 66.4%; Pred. No. 3.5e-39;  
Matches 233; Conservative 0; Mismatches 118; Indels 0; Gaps 0;

OY 59 atggaagatcactctgtgtgaaagagctcataccacaagagtcctgaccacaagag 118  
DB 148 ATGTGAGAGCTCTTGTGTGAGAAATGGGCTGAAAAAGGCCATGGATCTCTGAA 207  
OY 119 gaagatcaagcccatcaactacacccgtgcatggtgaaagcgtggtccctc 178  
DB 208 GAGATCAGATTCATCTCTTCTTCAAACTAATGSCCAATGCACTGGCAGCCCTT 267  
OY 179 cccaaagctggtggtcgttgaatggtgtaagagttgcagaaatgaatgaataactac 238  
DB 268 CCCAAAGCGCTGAGCTATGTAGATGCGGGAAGAGTTGCACTGGCGTGGAGAAATAT 327  
OY 239 ttgagcctgactcttaagaagaatttcaactgaagaagaatgagttatcatcaag 298  
DB 328 TTGGGACCATATTAAGAGGGGAATTTTACCAAGGAAGAAACAAATATATCAG 387  
OY 299 ctacacagttacttggaacaatggtcattgattgctggaagaattaccaggaagaca 358  
DB 388 TTAATGTAATGCTTGCAATATGATGCTGCAATATGCAAAATTTACAGAGCAACA 447  
OY 359 gataatgataaagaactactggaacacacacatcaaaagaagattata 409  
DB 448 GACATGAAATATAAATGTTTGGCACACCACTTGAGAGAAAGTCAAA 498

RESULT 2  
US-08-997-251-1  
Sequence 1, Application US/08997251  
Patent No. 6271440

GENERAL INFORMATION:

APPLICANT: GUBLER, FRANZ J.  
APPLICANT: JACOBSEN, JOHN V.  
TITLE OF INVENTION: PLANT REGULATORY PROTEINS III  
NUMBER OF SEQUENCES: 17  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Greenlee, Winner and Sullivan, P.C.  
STREET: 5370 Manhattan Boulevard  
CITY: Boulder  
STATE: CO  
COUNTRY: US  
ZIP: 80303

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/997,251  
FILING DATE: 23-DEC-1997  
CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: WO AU96/00383  
FILING DATE: 21-JUN-1996  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: AU PN6470/95  
FILING DATE: 09-NOV-1995  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: AU PN3779/95  
FILING DATE: 23-JUN-1995

ATTORNEY/AGENT INFORMATION:

NAME: Winner, Ellen P.  
REGISTRATION NUMBER: 28,547  
REFERENCE/DOCKET NUMBER: 110-97  
TELECOMMUNICATION INFORMATION:

TELEPHONE: (303) 499-8080  
TELEFAX: (303) 499-8089  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 2220 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: cDNA  
ORIGINAL SOURCE:

ORGANISM: Hordeum vulgare

STRAIN: Himalaya

DEVELOPMENTAL STAGE: Seed

TISSUE TYPE: Aleurone layer

IMMEDIATE SOURCE:

CLONE: HvGAMYB

FEATURE:

NAME/KEY: CDS

LOCATION: 275...1933

US-08-997-251-1

Query Match 10.9%; Score 109.6; DB 4; Length 2220;  
Best Local Similarity 59.2%; Pred. No. 3.2e-23;  
Matches 187; Conservative 0; Mismatches 129; Indels 0; Gaps 0;

OY 85 ggtctatccacaagaagtgctggaacaaagaagaatcaagcctcatcaactacat 144  
DB 391 GCCGCTCTGAAGAAAGGGCCCTGGACGTCCGGGAGAGCCATCTGTGGACTACTGT 450  
OY 145 ccgtgtccatggtgaagcgtctggtccctccccaagcgtggtgcttaagt 204  
DB 451 GAAGAGCAGCGCGGAGGAGACACGAGGAGGAGAGAGAGAGAGAGAGAGAGAGAG 510  
OY 205 tggtaagagtgcaagattgaatgataaactctggaagcctgatacttaagaagaaa 264  
DB 511 CGGCAAGAGCTCCGCTCCGCTGGGAGACACCTCAGGCCCAACCTCAAGAGAGGCG 570  
OY 265 ttcaactgaagaagaatgagcttatactcaagcttcaagcttacttgaagaagaatg 324  
DB 571 CTTACCCCCCGAGAGAGAGAGGCTCATCATCTCCTCAAGATGGCCACACTG 630  
OY 325 gtcattgattgctggaattaccagaagaagaagataatgataagaactactgga 384  
DB 631 GCGTGGAGAGCGCCGCTATTTGCCAGGCGCTACTGATGAATGAATGAATGAATGA 690  
OY 385 caccacacacaaaga 400  
DB 691 CACTCGAATTAAGAGA 706

RESULT 3  
US-08-997-251-3  
Sequence 3, Application US/08997251  
Patent No. 6271440

GENERAL INFORMATION:

APPLICANT: GUBLER, FRANZ J.  
APPLICANT: JACOBSEN, JOHN V.  
TITLE OF INVENTION: PLANT REGULATORY PROTEINS III  
NUMBER OF SEQUENCES: 17  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Greenlee, Winner and Sullivan, P.C.  
STREET: 5370 Manhattan Boulevard  
CITY: Boulder  
STATE: CO  
COUNTRY: US  
ZIP: 80303

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/997,251  
FILING DATE: 23-DEC-1997  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: WO AU96/00383  
FILING DATE: 21-JUN-1996  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: AU PN6470/95  
FILING DATE: 09-NOV-1995  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: AU PN3779/95  
FILING DATE: 23-JUN-1995  
ATTORNEY/AGENT INFORMATION:  
NAME: Winner, Ellen P.  
REGISTRATION NUMBER: 28,547  
REFERENCE/DOCKET NUMBER: 110-97  
TELEPHONE: (303) 499-8089  
TELEFAX: (303) 499-8080  
INFORMATION FOR SEQ ID NO: 3:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 2352 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: CDNA  
ORIGINAL SOURCE:  
ORGANISM: Oryza sativa  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 396..2054  
US-08-997-251-3

Query Match 10.9%; Score 109.2; DB 4; Length 2352;  
Best Local Similarity 59.8%; Pred. No. 4.3e-23;  
Matches 183; Conservative 0; Mismatches 123; Indels 0; Gaps 0;

QY 95 aacaaagtcgtgacaaagaagaagtcacagctcaccacacacccgttcacat 154  
DB 516 AAGAGGGGCGCATGACGCTCGGCGAGGAGCGCCATCTGTGAGCTACGTGAAGAACAC 575  
QY 155 ggtgaagctgctgagctccctccccaagctgctgagctgctgagctgctgaagt 214  
DB 576 GGGCAGGGGAGACTGGAAGCGGCTGCAGAGAACACCGGGCTGTCCGCTCGGAGAGAC 635  
QY 215 tgcagattgaatgataaactactgagcctgacatcctaagaaggaataatcactgaa 274  
DB 636 TGCCGCCCTCCGCTGGCGACACCTGAGGCCCAACCTCAAGAGGGGCGCTTCACCGCC 695  
QY 275 gaagaagctgagctatatacagcttcacagcttacttggaagaataatgctcattgac 334  
DB 696 GAGGAGGAGAGCTCATCATCCAGCTCCACTCCACAGATGGGGAACAGTGGCTCGGATG 755  
QY 335 gctgaagattaccagaagaacagataatgagataaagaactctggaacacacacatc 394  
DB 756 GCCCGTATTGGCCAGGGGCGCACTGATATGAATAAGAAATTACTGCAATCTGATA 815  
QY 395 aaaga 400  
DB 816 AAGAGA 821

RESULT 4  
PCT-US93-06251-90  
Sequence 90, Application PC/TUS9306251  
GENERAL INFORMATION:  
APPLICANT: Wickstrom, Eric and Rife, Jason P.  
TITLE OF INVENTION: Trivalent synthesis of Oligonucleotides containing  
NUMBER OF SEQUENCES: 93  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: SCULLY, SCOTT, MURPHY & PRESSER

STREET: 400 Garden City Plaza  
CITY: Garden City  
STATE: NY  
COUNTRY: USA  
ZIP: 11530  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: PCT/US93/06251  
FILING DATE: 19930630  
CLASSIFICATION:  
ATTORNEY/AGENT INFORMATION:  
NAME: Digilio, Frank S.  
REGISTRATION NUMBER: 31,345  
REFERENCE/DOCKET NUMBER: 8586  
TELEPHONE: 516-742-4343  
TELEFAX: 516-742-4366  
TELEX: 230 901 SAMS UR  
INFORMATION FOR SEQ ID NO: 90:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1035 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
PCT-US93-06251-90

Query Match 7.9%; Score 79.2; DB 5; Length 1035;  
Best Local Similarity 54.9%; Pred. No. 3e-14;  
Matches 178; Conservative 0; Mismatches 143; Indels 3; Gaps 1;

QY 82 aaagctctaccacaagaagctgctgacacaagaagaatcaagcctcacaacta 141  
DB 120 AACCCTGAGCTCATCATGAAGGCTCTTGACCAAGAAAGATCAGAGATGATAGACT 179  
QY 142 catcgctgcacatgctggaagctgctgagcttccctcccaagctgctgctgcttag 201  
DB 180 TGTACGAATAATGCGTCCGCAACAGTTGCTGTATTGCCAAGCACTTAAGGGGG--AG 236  
QY 202 agtgtgaagattgagattgaatgaatgaataaactactgagctgactcctaagaagag 261  
DB 237 AATTGAAAACAAATGTAGGAGAGGTGTCATACCACTTGATCCAGATCAGAAATTAAGAAAC 296  
QY 262 aaattcactgaagaagaatgagcttaccatcaagcttcacagcttacttggaacaa 321  
DB 297 CTTCTGAGAGAGAGAGAGAGAGATTTATTACAGACACACAGAGACTGGGGAACAG 356  
QY 322 atgttcattgattgctggaagattaccagaagaacagataatgagataaagaactactg 381  
DB 357 AGGGGAGAGAAATCGCAAGCTACTGCTGAGCAGAACTGATTAATCTATCAAGAACACTG 416  
QY 382 gaacacacacatcaaaaagaagct 405  
DB 417 GAATTTACAAATGCTCGGAAGGT 440

RESULT 5  
US-08-306-691B-45  
Sequence 45, Application US/08306691B  
Patent No. 5734039  
GENERAL INFORMATION:  
APPLICANT: Calabretta, Bruno  
APPLICANT: Skoraki, Tomasz  
TITLE OF INVENTION: OLIGONUCLEOTIDES TARGETING COOPERATING ONCOGENES  
NUMBER OF SEQUENCES: 55  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Seidel, Gonda, Lavorina & Monaco, P.C.



STATE: NY  
COUNTRY: USA  
ZIP: 11530  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: PCT/US93/06251  
FILING DATE: 19930630  
CLASSIFICATION:  
ATTORNEY/AGENT INFORMATION:  
NAME: DIGILIO, Frank S.  
REGISTRATION NUMBER: 31,346  
REFERENCE/DOCKET NUMBER: 8586  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 516-742-4343  
TELEFAX: 516-742-4366  
TELEX: 230 901 SANS UR  
INFORMATION FOR SEQ ID NO: 89:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 3230 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
PCT-US93-06251-89

Query Match  
Best Local Similarity 7.9%; Score 79.2; DB 5; Length 3230;  
Matches 178; Conservative 0; Mismatches 143; Indels 3; Gaps 1;

QY 82 aaagctatatacaacaagaagtcgagcgaagaagaagatcaacgctcacaacta 141  
DB 261 AAACCTGAGCTCATCAAGAGTCTGAGCAAGAGAGATGAGAGAGAGAGAGAGT 320  
QY 142 catcggttcacgttggaagtcgctgctgctcccccagaagctcgtgctgctag 201  
DB 321 TGTACAGAAATCGGCGCAAAAGCTGTTATGTCAGAGACCTTAAGGGG--AG 377  
QY 202 atgtgtaagagtgatgaatgaatgaatgaatgaatgaatgaatgaatgaatga 261  
DB 378 AATTGAAACAAATGATGAGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 437  
QY 262 aaattcactgaagaagaatgaatgaatgaatgaatgaatgaatgaatgaatga 321  
DB 438 CTCCGAGACAGAGAGAGAGAGAGATTTATTTACAGGACACAGAGAGAGAGAGAG 497  
QY 322 atgtgcatgtatgctggaagattacacgaagaagaagaagaagaagaagaagaaga 381  
DB 498 ATGGCAGAGAAATCGCAAACTGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 557  
QY 382 gaac 405  
DB 558 GAATCTACAAATGCGTCGGAAGGT 581

RESULT 8  
US-09-008-979A-6  
Sequence 6, Application US/09008979A  
Patent No. 6080914  
GENERAL INFORMATION:  
APPLICANT: Conner, Timothy W.  
TITLE OF INVENTION: Strawberry Promoters and Genes  
NUMBER OF SEQUENCES: 14  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Carmen Rodriguez, Paralegal, Arnold, White & Durkee  
STREET: P.O. Box 4433  
CITY: Houston  
STATE: Texas  
COUNTRY: USA

ZIP: 77210-4433  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/008,979A  
FILING DATE:  
CLASSIFICATION: 800  
ATTORNEY/AGENT INFORMATION:  
NAME: Kammerer, Patricia A.  
REGISTRATION NUMBER: 29,775  
REFERENCE/DOCKET NUMBER: MOBT:058 (38-21(10613)A)  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (713) 787-1438  
TELEFAX: (713) 789-2679  
INFORMATION FOR SEQ ID NO: 6:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 785 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
US-09-008-979A-6

Query Match  
Best Local Similarity 6.9%; Score 69.8; DB 3; Length 785;  
Matches 107; Conservative 0; Mismatches 42; Indels 2; Gaps 1;

QY 281 gatgagcttcatcaacagctcacaagttacttggaacaacatgctgattgctgga 340  
DB 2 GTTGATTGATCATCAGGCTTCATAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 61  
QY 341 agattccaggaagaacagataatgagataagaactctggaacacacacacacacac 400  
DB 62 CGATGCGAG 119  
QY 401 aagctataagcagagagagatgattcaccacaa 431  
DB 120 AAAAGATCAAAAGAGCGGCTTCATACGCAAA 150

RESULT 9  
US-09-460-618-6  
Sequence 6, Application US/09460618  
Patent No. 6235482  
GENERAL INFORMATION:  
APPLICANT: Conner, Timothy W.  
TITLE OF INVENTION: Strawberry Promoters and Genes  
NUMBER OF SEQUENCES: 14  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Carmen Rodriguez, Paralegal, Arnold, White & Durkee  
STREET: P.O. Box 4433  
CITY: Houston  
STATE: Texas  
COUNTRY: USA  
ZIP: 77210-4433  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/460,618  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/09/008,979  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Kammerer, Patricia A.  
REGISTRATION NUMBER: 29,775



Query match	6.3%; Score 63.6; DB 3; Length 469;
-------------	-------------------------------------

LOCATION: 2669..26/4

OTHER INFORMATION: /label- SnabI  
FEATURE:  
NAME/KEY: -  
LOCATION: 2634..2639  
OTHER INFORMATION: /label- SnabI  
FEATURE:  
NAME/KEY: -  
LOCATION: 3008..3013  
OTHER INFORMATION: /label- HpaI  
FEATURE:  
NAME/KEY: -  
LOCATION: 1..1077  
OTHER INFORMATION: /label- Pci  
OTHER INFORMATION: /note- "region containing promoter of C1 gene"  
FEATURE:  
NAME/KEY: -  
LOCATION: 1078..2134  
OTHER INFORMATION: /label- C1  
OTHER INFORMATION: /note- "coding region of C1 gene"  
FEATURE:  
NAME/KEY: -  
LOCATION: 2135..2430  
OTHER INFORMATION: /label- 3'C1  
OTHER INFORMATION: /note- "region containing polyadenylation signal"  
FEATURE:  
NAME/KEY: -  
LOCATION: 1033..1038  
OTHER INFORMATION: /label- TATA-Box  
FEATURE:  
NAME/KEY: -  
LOCATION: 1061..1062  
OTHER INFORMATION: /label- transcript-init  
OTHER INFORMATION: /note- "transcription initiation site"  
FEATURE:  
NAME/KEY: -  
LOCATION: 1211..1299  
OTHER INFORMATION: /label- 1299  
FEATURE:  
NAME/KEY: -  
LOCATION: 1430..1575  
OTHER INFORMATION: /label- 1575  
FEATURE:  
NAME/KEY: -  
LOCATION: 935..939  
OTHER INFORMATION: /label- C1-S  
OTHER INFORMATION: /note- "TGCAG sequence (in C1 gene) which in the  
OTHER INFORMATION: C1-S sequence is changed to TTAGC"  
US-08-485-139-1

Query Match 5.6%; Score 56; DB 2; Length 4059;  
Best Local Similarity 63.2%; Pred. No. 6.1e-07;  
Matches 86; Conservative 0; Mismatches 50; Indels 0; Gaps 0;  
QY 185 gctgctggtctgctagatggtgtaagagtgacagtaagatgataaactctgag 244  
Db 1293 -ggtcagaggttgcgtgcgtgcgcgacagctccgctgcgtgcgtgacatcctccg 1352  
QY 245 cctgactcttaagagaggaattctcctgaagaagaagatgagcttcatcaagcttcac 304  
Db 1353 cccacacatcagagcggcggaacatctcctacgacagagagatctcatcaccctccac 1412  
QY 305 agttacttgaaaca 320  
Db 1413 AGGCTCTCGGCACACA 1428

RESULT 14  
US-08-750-357-1  
Sequence 1, Application US/08750357  
Patent No. 6008437  
GENERAL INFORMATION:  
APPLICANT: KREBBERS, ENNO  
APPLICANT: WILLIAMS, MARK

APPLICANT: LEEMANS, Jan  
TITLE OF INVENTION: USE OF ANTHOCYANIN GENES TO MAINTAIN  
TITLE OF INVENTION: MALE STERILE PLANTS  
NUMBER OF SEQUENCES: 8  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: BURNS, DOANE, SWECKER & MATHIS, L.L.P.  
STREET: P.O. Box 1404  
CITY: Alexandria  
STATE: Virginia  
COUNTRY: United States  
ZIP: 22313-1404  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/750,357  
FILING DATE: 21-MAR-1997  
CLASSIFICATION: 800  
ATTORNEY/AGENT INFORMATION:  
NAME: McGowan, Malcolm K.  
REGISTRATION NUMBER: 39,300  
REFERENCE/DOCKET NUMBER: 018030-100  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (703) 836-6620  
TELEFAX: (703) 836-2021  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 4059 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
ORIGINAL SOURCE:  
ORGANISM: C1 gene of Zea mays  
FEATURE:  
NAME/KEY: -  
LOCATION: 279..284  
OTHER INFORMATION: /label- HpaI  
FEATURE:  
NAME/KEY: -  
LOCATION: 447..452  
OTHER INFORMATION: /label- EcoRI  
FEATURE:  
NAME/KEY: -  
LOCATION: 1735..1740  
OTHER INFORMATION: /label- AatII  
FEATURE:  
NAME/KEY: -  
LOCATION: 1505..1510  
OTHER INFORMATION: /label- EcoRI  
FEATURE:  
NAME/KEY: -  
LOCATION: 2081..2086  
OTHER INFORMATION: /label- XhoI  
FEATURE:  
NAME/KEY: -  
LOCATION: 2418..2430  
OTHER INFORMATION: /label- SfiI  
FEATURE:  
NAME/KEY: -  
LOCATION: 2669..2674  
OTHER INFORMATION: /label- SnabI  
FEATURE:  
NAME/KEY: -  
LOCATION: 2634..2639  
OTHER INFORMATION: /label- SnabI  
FEATURE:  
NAME/KEY: -  
LOCATION: 3008..3013  
OTHER INFORMATION: /label- HpaI  
FEATURE:

```
NAME/KEY: -
LOCATION: 1..1077
OTHER INFORMATION: /label= PCl
OTHER INFORMATION: /note= "region containing promoter of C1 gene"
FEATURE:
NAME/KEY: -
LOCATION: 1078..2134
OTHER INFORMATION: /label= C1
OTHER INFORMATION: /note= "coding region of C1 gene"
FEATURE:
NAME/KEY: -
LOCATION: 2135..2430
OTHER INFORMATION: /label= 3' C1
OTHER INFORMATION: /note= "region containing polyadenylation signal of C1 gene"
FEATURE:
NAME/KEY: -
LOCATION: 1033..1038
OTHER INFORMATION: /label= TATA-Box
FEATURE:
NAME/KEY: -
LOCATION: 1061..1062
OTHER INFORMATION: /label= transcript-init
OTHER INFORMATION: /note= "transcription initiation site"
FEATURE:
NAME/KEY: -
LOCATION: 1211..1299
FEATURE:
NAME/KEY: -
LOCATION: 1430..1575
FEATURE:
NAME/KEY: -
LOCATION: 935..939
OTHER INFORMATION: /label= C1-S
OTHER INFORMATION: /note= "TGCAG sequence (in C1 gene) which in the C1-S sequence"
OTHER INFORMATION: /note= "is changed to TTAGG"
US-08-750-357-1

Query Match
Best Local Similarity 5.6%; Score 56; DB 3; Length 4059;
Matches 86; Conservative 0; Mismatches 50; Indels 0; Gaps 0;

QY 185 gctgtggtggtccttagatgtgtgaagattgagatgaagataaactacttgaag 244
DB 1293 GGTGAGGTTTGGCTGGTGGCGCAAGAGCTGCCGCTGGCTGGAATCTACCTCCGG 1352

QY 245 cctgacttaagaagaaatttcactgaagaagaagatgagctatcatcaagcttcac 304
DB 1353 CCCAACAATCAGCGCGGCAACATCTCTACGACGAGGAGGATCTCATCATCTCCGCTCCAC 1412

QY 305 agttacttggaaaca 320
DB 1413 AGGCTCTCGGCACA 1428

RESULT 15
US-08-485-139-5
Sequence 5, Application US/08485139
Patent No. 5880331
GENERAL INFORMATION:
APPLICANT: KREBBERS, ENO
APPLICANT: WILLIAMS, MARK
TITLE OF INVENTION: USE OF ANTHOCYANIN GENES TO MAINTAIN
TITLE OF INVENTION: MALE STERILE PLANTS
NUMBER OF SEQUENCES: 6
CORRESPONDENCE ADDRESS:
ADDRESSEE: Burns, Doane, Swecker & Mathis
STREET: P. O. Box 1404
CITY: Alexandria
STATE: Virginia
COUNTRY: United States
ZIP: 22313-1404
```

```
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/485,139
FILING DATE: 07-JUN-1995
CLASSIFICATION: 800
ATTORNEY/AGENT INFORMATION:
NAME: McGowan, Malcolm K.
REGISTRATION NUMBER: 39,300
REFERENCE/DOCKET NUMBER: 010830-096
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 836-6620
TELEFAX: (703) 836-2021
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 4824 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: circular
MOLECULE TYPE: DNA (genomic)
ORIGINAL SOURCE:
ORGANISM: plasmid pCOL9
FEATURE:
NAME/KEY: -
LOCATION: 396..401
OTHER INFORMATION: /label= EcoRI
FEATURE:
NAME/KEY: -
LOCATION: 2367..2379
OTHER INFORMATION: /label= SfiI
FEATURE:
NAME/KEY: -
LOCATION: 884..888
OTHER INFORMATION: /label= C1-S
OTHER INFORMATION: /note= "TGCAG (in C1) which in C1-S allele is
OTHER INFORMATION: replaced with TTAGG"
US-08-485-139-5
```

```
Query Match
Best Local Similarity 5.6%; Score 56; DB 2; Length 4824;
Matches 86; Conservative 0; Mismatches 50; Indels 0; Gaps 0;

QY 185 gctgtggtggtccttagatgtgtgaagattgagatgaagataaactacttgaag 244
DB 1242 GGTGAGGTTTGGCTGGTGGCGCAAGAGCTGCCGCTGGCTGGAATCTACCTCCGG 1301

QY 245 cctgacttaagaagaaatttcactgaagaagaagatgagctatcatcaagcttcac 304
DB 1302 CCCAACAATCAGCGCGGCAACATCTCTACGACGAGGAGGATCTCATCATCTCCGCTCCAC 1361

QY 305 agttacttggaaaca 320
DB 1362 AGGCTCTCGGCACA 1377
```

Search completed: January 23, 2002, 20:22:30  
Job time: 5633 sec



GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

## OM nucleic - nucleic search, using sw model

Run on: January 23, 2002, 18:40:17 ; Search time 1542.92 Seconds  
(without alignments)  
7006.360 Million cell updates/sec

Title: US-09-453-387A-1

Perfect score: 1006  
Sequence: 1 taacacgcttattcttc.....aatgcttaattataaaaa 1006

Scoring table: IDENTITY\_MUC  
Gapop 10.0 , Gapext 1.0

Searched: 11351937 seqs, 5372889281 residues

Total number of hits satisfying chosen parameters: 22703874

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

## Database :

EST:\*  
1: em\_estfun:\*  
2: em\_esthm:\*  
3: em\_estlin:\*  
4: em\_estom:\*  
5: em\_estopl:\*  
6: em\_estlda:\*  
7: em\_estro:\*  
8: em\_estrov:\*  
9: em\_hlc:\*  
10: gb\_est1:\*  
11: gb\_est2:\*  
12: gb\_hlc:\*  
13: gb\_gss:\*  
14: em\_gss\_fun:\*  
15: em\_gss\_hum:\*  
16: em\_gss\_inv:\*  
17: em\_gss\_pln:\*  
18: em\_gss\_pro:\*  
19: em\_gss\_rnd:\*  
20: em\_gss\_vrt:\*  
21: em\_gss\_other:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	533.4	53.0	857	11	BF269948 GA_EB000
2	361	35.9	927	11	BG445164 GA_EB002
3	295.8	29.4	757	11	BG889982 EST515833
4	288.6	28.7	526	11	BE801303 sr12d04.y
5	287.4	28.6	562	10	A1771837 EST252937
6	283	28.1	447	10	A1486301 EST244622
7	278	27.6	536	10	BE124666 EST393701
8	275	27.3	549	11	BE802821 sr45a11.y
9	261.4	26.0	603	10	A1731905 BNLGH1112
10	261	25.9	911	11	BG442984 GA_EB001
11	256.8	25.5	730	10	AM775893 EST334958
12	253.4	25.4	603	11	BF716393 sr119f01

13	254.2	25.3	580	10	A1897672
14	247.4	24.6	468	10	AL373528
15	240.2	23.9	545	11	BE643814
16	239.2	23.8	524	11	BE596995
17	236.4	23.5	494	11	BE517838
18	235	23.4	457	10	AU082322
19	235	23.4	467	10	AU082325
20	235	23.4	492	10	AM279057
21	235	23.4	510	11	BE804284
22	233.4	23.2	501	11	BG237705
23	233.4	23.2	522	10	BE331171
24	233	23.2	385	11	BE805745
25	232.4	23.1	683	10	BE052585
26	231.2	23.0	425	10	AV408163
27	230.6	22.9	434	10	A1736697
28	226.2	22.5	704	11	BI421740
29	225.8	22.4	446	10	AU184006
30	225.4	22.4	445	10	AU173093
31	220.6	21.9	385	11	BE804943
32	219.6	21.8	564	10	BE611947
33	219.6	21.8	572	10	AM737657
34	219.2	21.8	579	10	A1727146
35	219.2	21.8	620	10	AM011253
36	218.4	21.7	578	10	A1730649
37	217.6	21.6	617	10	BG301049
38	217.6	21.6	624	11	BG301022
39	217.2	21.6	625	11	AM064634
40	216.6	21.5	626	11	BE519041
41	214.8	21.4	656	11	T42000
42	213.4	21.2	473	11	BF627287
43	207.2	20.6	756	11	AU183276
44	205.6	20.4	451	10	BE361115
45	201.6	20.0	590	10	

## ALIGNMENTS

RESULT 1  
LOCUS BF269948 857 bp mRNA EST 07-MAR-2001  
DEFINITION GA\_EB0006B13f Gossypium arboreum 7-10 dpa fiber library Gossypium  
arboireum CDNA clone GA\_EB0006B13f, mRNA sequence.

ACCESSION BF269948  
VERSION BF269948.1 GI:11200943

KEYWORDS EST.  
SOURCE Gossypium arboreum.

## ORGANISM

Gossypium arboreum.  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
Rosidae; eustosids II; Malvales; Malvaceae; Gossypium.

REFERENCE 1 (bases 1 to 857)  
Wing,R.A., Frisch,D., Yu,Y., Main,D., Rambo,T., Simmons,J., Henry  
,D., Wood,T.C., Leslie,A. and Wilkins,T.A.

An integrated analysis of the genetics, development, and evolution  
of the cotton fiber  
Unpublished (2000)

JOURNAL CONTACT: Wing RA  
Clemson University Genomics Institute  
Clemson University  
100 Jordan Hall, Clemson, SC 29634, USA

TELEPHONE: Tel: 864 656 7288  
FAX: Fax: 864 656 4293  
EMAIL: Email: rwing@clemson.edu

FEATURES  
Seq primer: TAATACGACTGCTATAGGG  
High quality sequence start: 4  
High quality sequence stop: 633.  
Location/Qualifiers

1..857  
/organism="Gossypium arboreum"  
/strain="AKA"  
/cdliver="8400"  
/db\_xref="taxon:29723"



Db	244	TCAACTTCATAGATTGCTTGGAAACAAATGGTCAAGTTATAGCTGGAGGATTACTCGGAA	303
QY	354	gaacagataatgagataaagaactacttgaacacacacatcaaaagaagctttaagca	413
Db	304	GAATGTATAGAAAGAAATCAAGAACTATTGGAATATACATATTTAAACGAAAGCTCATCACTC	363
QY	414	gaggaaatgatccacaacactatcgtctctccaaatcaaacggcgacata---ccaaagcag	470
Db	364	GTGGCATTTAGTCTTCAAACTCAACGCTCCACTCTACACACACCGCCACTACCTCCACACCA	423
QY	471	tcacagccccacacgaattggtatttcagaaaatcgccacatcgtttccaaatcagtc	530
Db	424	CCACCATCAACACCGCAGTCAAAACCCACACACACACACACCAAAAACATCAACATGG	483
QY	531	ccatcaaaaac 541	
Db	484	ATTTCACAAAC 494	
RESULT	4		
LOCUS	BE801303	526 bp mRNA	21-NOV-2000
DEFINITION	sr12d04.y1 Gm-cl050 Glycine max cDNA clone GENOME SYSTEMS CLONE ID: Gm-cl050-128 5' similar to TR:P81395 P81395 MYB-RELATED PROTEIN 330..17, mRNA sequence.		
ACCESSION	BE801303		
VERSION	BE801303.1	GI:10232415	
KEYWORDS	EST.		
SOURCE	soybean.		
ORGANISM	Glycine max		
REFERENCE	Eukaryotes; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eustosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae; Glycine.		
AUTHORS	1 (bases 1 to 526)		
JOURNAL	Shoemaker, R., Kelm, P., Vodkin, L., Expelling, J., Coryell, V., A., Bolla, B., Marra, M., Hillier, L., Kucaba, T., Martin, J., Wylie, T., Underwood, K., Steptoe, M., Theising, B., Allen, M., Y., Person, B., Swaller, T., Gibbons, M., Pape, D., Harvey, N., R., Ritter, E., Kohn, S., Shin, T., Jackson, Y., Cardenas, M., R., Waterston, R. and Wilson, R.		
COMMENT	Public Soybean EST Project Unpublished (1999) Contact: Shoemaker R./Public Soybean EST Project Public Soybean EST Project Washington University School of Medicine 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA Tel: 314 286 1800 Fax: 314 286 1810 Email: est@watson.wustl.edu This clone is available through: Genome Systems, Inc. 4633 World Parkway Clirke St. Louis, Missouri 63134 For further information call: (800) 430-0030 or (314) 427-3222 FAX: (888) 919-3324 or (314) 427-3324 or contact: clones@genomesystems.com or info@genomesystems.com web site: www.genomesystems.com Insert length: 1223 Std Error: 0.00 High quality sequence stop: 470. Location/Qualifiers 1..526 /organism="Glycine max" /db_xref="taxon:3847" /clone="GENOME SYSTEMS CLONE ID: Gm-cl050-128" /clone_11b="Gm-cl050" /tissue_type="leaf tissue at various developmental stages of greenhouse grown plants" /dev_stage="3 week old" /lab_host="DH10B" /note="Vector: pBluescript II SK+, Site_1: EcoRI; Site_2: XhoI; The Clark N1L was constructed and seed was provided by Dr. J. Specht, University of Nebraska (Shoemaker and Specht, 1995). The cDNA library was constructed from RNA isolated from leaf tissue at various developmental stages		





RESULT	7
BE124666	
LOCUS	BE124666 536 bp
DEFINITION	EST393701 GVN Medicago truncatula cDNA clone
ACCESSION	sequence. BE124666
	EST 07-SEP-2000
	pcvtn 6746, mRNA

Query Match	27.6%	Score 278;	DB 10;	Length 536;
Best Local Similarity	78.7%	Pred. No. 5, 1e-50;		
Matches 332;	Conservative 0;	Mismatches 90;	Indels 0;	Gaps 0;
QY 59	atggagcatcaacctgtgtgtgaaaaggctctataccacaagaagtgcctgagcccaagaag			118
Db 27	ATGGGCAGATCTCCTGTTGTGAGAAAGAGCACAAACAAAGAGTCTGGCAAAAGAA			86
QY 119	gaagatcaagccctcatctaactacatccgtgtcatagtgtgaagctgtctgagcttcctc			178
Db 87	GAAAGTGAACGCTCTCATTAACATCAAGCTCTCATGGGGAAGGTTGTTGGAGGCTCTT			146
QY 179	cccaaaagctgtcggcgctgcttagatgtgtgtanagattgcagaattaaagttgataaacac			238
Db 147	CCAAAGCTGCTGCTGCTTGCCTAAGATGTGGCAAGAGTTGCGAGATTGAGATTGATAATTAC			206
QY 239	ttgagcgctgtatccttaagaagagaaatttcactgaagaagaagatgagcttatcatcaag			298
Db 207	CTAAGACCTGATCTCTCAAGAGAGAAACTTCATGAACAAGAAAGATGACTTATCTATTAC			266
QY 299	cttcacagtttaactcttggaanaacaatagtcattgtattgtctgaagattccacggaagaaca			358

Db 267 CTCACACCTACTTGGAACAATGCTCTTAATAGCTGATGATTACCGAGNAGACA 326  
 Oy 359 gataatgataaagaactactggaacacacacatcaaaagaagcttataagcagagga 418  
 Db 327 GATAATGAGATCAAGAAATTACTGGAACACTCATCAAGGAAAGCTCTACAGCGTGA 386  
 Oy 419 atgatccacaacactactgctctcaatcaaacgagccataccaacagtcagagcc 478  
 Db 387 GTCGACCTCAACACATCGCTACTTAACGACTCAACGATCAACACCATCAATATCCA 446  
 Oy 479 cc 480  
 Db 447 CC 448

RESULT 8  
 BE802821 549 bp mRNA EST 21-NOV-2000  
 LOCUS sr45a11.y1 Gm-cl051 glycine max cDNA clone GENOME SYSTEMS CLONE ID:  
 DEFINITION Gm-cl051-981 5' similar to TR:P81395 P81395 MYB-RELATED PROTEIN  
 330 ; mRNA sequence.  
 ACCESSION BE802821  
 VERSION BE802821.1 GI:10233933  
 KEYWORDS soybean.  
 SOURCE Glycine max  
 ORGANISM Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
 Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae;  
 Glycine.  
 1 (bases 1 to 549)  
 REFERENCE Shoemaker, R., Kelm, P., Vodkin, L., Erpelidg, J., Coryell, V., Khanna  
 A., Bolle, B., Maritz, M., Hillier, L., Kucaba, T., Martin, J., Beck, C.,  
 Wylie, T., Underwood, K., Steptoe, M., Theising, B., Allen, M., Bowers  
 , Y., Peterson, B., Swaller, T., Gibbons, M., Pape, D., Harvey, N., Schurk  
 , R., Ritters, E., Kohn, S., Shih, T., Jackson, Y., Cardenas, M., McCann  
 , R., Waterston, R., and Wilson, R.  
 TITLE Public Soybean EST Project  
 JOURNAL Unpublished (1999)  
 COMMENT Contact: Shoemaker R/Public Soybean EST Project  
 Public Soybean EST Project  
 Washington University School of Medicine  
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA  
 Tel: 314 286 1800  
 Fax: 314 286 1810  
 Email: est@wustl.edu  
 This clone is available through: Genome Systems, Inc. 4633 World  
 Parkway Circle St. Louis, Missouri 63134 For further information  
 call: (800) 430-0030 or (314) 427-3222 FAX: (888) 919-3324 or (314)  
 427-3324 or contact: clones@genomesystems.com or  
 info@genomesystems.com web site: www.genomesystems.com  
 Insert length: 1111 Std Error: 0.00  
 High quality sequence stop: 471.  
 Location/Qualifiers  
 1. 549  
 /organism="Glycine max"  
 /db\_xref="taxon:3847"  
 /clone="GENOME SYSTEMS CLONE ID: Gm-cl051-981"  
 /clone\_1lb="Gm-cl051"  
 /tissue\_type="floral meristematic mRNA"  
 /lab\_host="DH10B"  
 /note="Vector: pBluescript II SK+; Site 1: EcoRI; Site 2:  
 XhoI; The cDNA library was constructed from floral  
 meristematic mRNA provided by Dr. Halina Knap of Clemson  
 University. Complementary DNA was synthesized from mRNA  
 using a primer consisting of a poly(dT) sequence with a  
 XhoI restriction site. EcoRI adapters were ligated to the  
 blunt-ended cDNA fragments followed by XhoI digestion. The  
 cDNA fragments were directionally cloned into the  
 EcoRI-XhoI restriction site of the pBluescript vector. The  
 ligated cDNA fragments were transformed into DH10B host  
 cells (GibcoBRL). This library was constructed in the  
 laboratory of Dr. Randy Shoemaker."

BASE COUNT 186 a 142 c 108 g 113 t  
 ORIGIN  
 Query Match 27.3%; Score 275; DB 11; Length 549;  
 Best Local Similarity 78.9%; Pred. No. 2.3e-49;  
 Matches 340; Conservative 0; Mismatches 90; Indels 1; Gaps 1;  
 Oy 54 aactatgtagagacatcacctctgtgtgaagaagcttaccacaaagaagtcgcctgacaa 113  
 Db 35 AATCCATGGAGATGCCCTGTGTGCGAGAAAGACACACCAACAAAGAGCTTGACCA 94  
 Oy 114 aagagaagatcaacagcgcctcatcaatcacatccgtgcacatgtgtaagagctgtgcgt 173  
 Db 95 AAGAGAAAGACGACGAGCCTCATCATCACTACATCAAGCTCCATGAGGAGGCTTGAGAT 154  
 Oy 174 cccctcccaaaagctgtgtgctgtccttagatgtgtgtaagaagttgcaagattgaagatgataa 233  
 Db 155 CCTCCCAAGAGCTGCTGCTGCTCAGATGTGGCAAGAGTTGACAGACTCAGATGATTA 214  
 Oy 234 actacttggagcctgtatcttaagaagagaaattcacactgaagaagaatgaagcttaca 293  
 Db 215 ATTACTCAGACCTGATC-TCAGAGAGGCAACTTCACTGAAGAGAGAGATGACTCATCA 273  
 Oy 294 tcaagctcacagatttacttctgtgaacacaaatggtcatgtatgtctgtgaagattaccaggaa 353  
 Db 274 TAACTCCTCATAGCTTACTTCTGGAACAATGCTTTTGATAGCTGCAAGGTTACTGTGAA 333  
 Oy 354 gaacagataatgataaagaactactgtgaacacacacatcaaaagaagcttataagca 413  
 Db 334 GAACGATACGAAATTAATAAATATGGAACACTCACTCAATCAAGAAAGAACTCTACACC 393  
 Oy 414 gaggatgatgcacaacactcgtctctcaatcaaaagcagcaatacaacagatca 473  
 Db 394 GCGGATCAACCTCTGACTCATGCTCACTCAACGCTGCTCCGCCCTCCGGCAACCA 453  
 Oy 474 cagcccccacc 484  
 Db 454 CCGTCACAGCC 464

RESULT 9  
 A1731905 603 bp mRNA EST 11-JUN-1999  
 LOCUS BNRGH11247 Six-day Cotton fiber Gossypium hirsutum cDNA 5' similar  
 DEFINITION to myb protein 308 - garden snapdragon, mRNA sequence.  
 ACCESSION A1731905  
 VERSION A1731905.1 GI:5050757  
 KEYWORDS EST.  
 SOURCE upland cotton.  
 ORGANISM Gossypium hirsutum  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
 Rosidae; eurosids II; Malvales; Malvaceae; Gossypium.  
 REFERENCE 1 (bases 1 to 603)  
 AUTHORS Blewitt, M., Matz, E.C., Davy, D.F. and Burr, B.  
 TITLE ESTs from developing cotton fiber  
 JOURNAL Unpublished (1999)  
 COMMENT Contact: Ben Burr  
 Biology Department  
 Brookhaven National Laboratory  
 Upton, NY 11973, USA  
 Tel: 516-344-3396  
 Fax: 516-344-3407  
 Email: burrb@bnl.bnl.gov  
 Seq primer: T3 primer.  
 Location/Qualifiers  
 1. 603  
 /organism="Gossypium hirsutum"  
 /cultivar="Acala Maxxa"  
 /db\_xref="taxon:3635"  
 /clone\_1lb="Six-day Cotton fiber"  
 /tissue\_type="immature fiber"



## COMMENT

Contact: Deborah A. Samac  
 Department of Plant Pathology  
 University of Minnesota  
 495 Borlaug Hall, 1991 Upper Buford Circle, St. Paul, MN 55108, USA  
 Tel: 612 625 1243  
 Fax: 651 649 5058  
 Email: debby@purdue1.crl.umn.edu  
 Minnesota sequence name: M259226e  
 TIGR sequence name: MTEA0821K  
 More information is available at:  
<http://chryslr.tamu.edu/medicago>  
 Seq. primer: SKmod (CTA GAA CTA gta gat cc).

## FEATURES

## SOURCE

1. 730

/organism="Medicago truncatula"

/cultivar="genotype A17"

/db\_xref="taxon:3880"

/clone="pDSIL-3M20"

/tissue="leaf"

/tissue\_type="Leaves infected with Colletotrichum trifolii"

/dev\_stage="cotyledons and primary leaves harvested 5 and 8 days after inoculation with Colletotrichum trifolii"

/lab\_host="E. coli strain XL0R"

/note="Vector: pBluescript SK-; Site\_1: EcoRI; Site\_2: XhoI; cDNA was prepared from polyA+ enriched RNA from cotyledons and primary leaves harvested 5 and 8 days after inoculation with Colletotrichum trifolii. The cDNA was directionally ligated into the Uni-ZAP XR vector from StrataGene and packaged using Gigapack III Gold packaging extracts. Plasmids containing lambda-ZAP inserts were excised from the recombinant lambda-ZAP phage using Ex-Aassist helper phage and propagated in XL0R cells. Note: EST may be of fungal origin."

BASE COUNT 278 a 140 c 119 g 193 t

## ORIGIN

Query Match 25.5%; Score 256.8; DB 10; Length 730;

Best Local Similarity 75.0%; Pred. No. 1.8e-45;

Matches 321; Conservative 0; Mismatches 107; Indels 0; Gaps 0;

## FEATURES

## SOURCE

1. 603

/organism="Glycine max"

/db\_xref="taxon:3847"

/clone="GENOME SYSTEMS CLONE ID: Gm-c1058-1801"

/tissue="leaf"

/tissue\_type="Hypocotyl"

/lab\_host="DH10B"

/note="Vector: pBluescript II SK-; Site\_1: EcoRI; Site\_2: XhoI; The cDNA library was constructed from mRNA isolated from hypocotyl tissue of 2 week old etiolated seedlings for pL468916. Complementary DNA was synthesized from mRNA using a primer consisting of a poly(dT) sequence with a XhoI restriction site. EcoRI adapters were ligated to the blunt-ended cDNA fragments followed by XhoI digestion. The cDNA fragments were directionally cloned into the EcoRI-XhoI restriction site of the pBluescript vector. The ligated cDNA fragments were transformed into DH10B host cells (Gibco BRL). This library was constructed in the laboratory of Dr. Randy Shoemaker at Iowa state university."

## BASE COUNT

189 a 149 c 133 g 132 t

## ORIGIN

Query Match 25.4%; Score 255.4; DB 11; Length 603;

Best Local Similarity 74.4%; Pred. No. 3.8e-45;

Matches 322; Conservative 0; Mismatches 111; Indels 0; Gaps 0;

## FEATURES

## SOURCE

1. 603

/organism="Glycine max"

/db\_xref="taxon:3847"

/clone="GENOME SYSTEMS CLONE ID: Gm-c1058-1801"

/tissue="leaf"

/tissue\_type="Hypocotyl"

/lab\_host="DH10B"

/note="Vector: pBluescript II SK-; Site\_1: EcoRI; Site\_2: XhoI; The cDNA library was constructed from mRNA isolated from hypocotyl tissue of 2 week old etiolated seedlings for pL468916. Complementary DNA was synthesized from mRNA using a primer consisting of a poly(dT) sequence with a XhoI restriction site. EcoRI adapters were ligated to the blunt-ended cDNA fragments followed by XhoI digestion. The cDNA fragments were directionally cloned into the EcoRI-XhoI restriction site of the pBluescript vector. The ligated cDNA fragments were transformed into DH10B host cells (Gibco BRL). This library was constructed in the laboratory of Dr. Randy Shoemaker at Iowa state university."

## RESULT 12

## LOCUS

## DEFINITION

## ACCESSION

## VERSION

## KEYWORDS

## SOURCE

## ORGANISM

## REFERENCE

## AUTHORS

## TITLE

## JOURNAL

## COMMENT

## REFERENCE

## AUTHORS

## TITLE

## JOURNAL

## COMMENT

## REFERENCE

## AUTHORS

## TITLE

## JOURNAL

## COMMENT

## REFERENCE

## AUTHORS

## TITLE

## JOURNAL

## COMMENT

## REFERENCE

## AUTHORS

## TITLE

## JOURNAL

## COMMENT

## REFERENCE

## AUTHORS

## TITLE

## JOURNAL

## COMMENT

## REFERENCE

## AUTHORS

## TITLE

## JOURNAL

## COMMENT

## REFERENCE

## AUTHORS

## TITLE

## JOURNAL

## COMMENT

## REFERENCE

## AUTHORS

## TITLE

## JOURNAL

## COMMENT

## REFERENCE

## AUTHORS

## TITLE

## JOURNAL

## COMMENT

## REFERENCE

## AUTHORS

## TITLE

## JOURNAL

## COMMENT

## REFERENCE

## AUTHORS

## TITLE

## JOURNAL

## COMMENT

## REFERENCE

## AUTHORS

## TITLE

## JOURNAL

## COMMENT

## REFERENCE

## AUTHORS

## TITLE

## JOURNAL

## COMMENT

## REFERENCE

## AUTHORS

## TITLE

## JOURNAL

## COMMENT



COMMENT

CONTACT: DIXON RA







Modulating transcription in plants for improving cotton fiber quality,  
comprises introducing a recombinant expression cassette comprising a  
promoter operably linked to a sequence encoding a MYB polypeptide -  
Claim 5; Page 41; 50pp; English.

The present sequence is that of cotton MYB transcription factor  
GhMYB 1. The protein includes a highly conserved N-terminal  
DNA-binding domain, and a non-conserved C-terminal transcriptional  
regulatory domain. It also contains a Cys-rich domain similar to a  
zinc finger motif. The sequence was deduced from GhMYB 1 cDNA  
isolated from a Gossypium hirsutum cv. Acala 5J-2 3-day pre-anthesis  
ovule cDNA library. GhMYB 1 transcripts are relatively abundant and  
appear to be expressed in all tissues examined. A claimed method  
of modulating transcription in a plant involves introducing into  
the plant a recombinant expression cassette comprising a promoter  
operably linked to a heterologous polynucleotide encoding a MYB  
protein, especially GhMYB 1 or GhMYB 6. The plant is especially a  
cotton plant, and the promoter directs expression of the  
polynucleotide in cotton fibres or roots. Valuable phenotypes  
can be conferred on the plant by this method, e.g. increased fibre  
quality, alteration of root architecture, and enhanced growth.  
The GhMYB 1 gene is particularly useful for increasing levels of  
gene expression in plant fibres to increase fibre yield, length,  
strength and fineness.

Sequence 293 AA:

Query Match 99.0%; Score 1567.5; DB 22; Length 293;  
Best Local Similarity 99.7%; Pred. No. 5.2e-141;  
Matches 293; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

1 MGRSPCEKAHNTKGAATWTEEDRLNIVYHGGCWRSLPKAGLRLCGSCRLRWNY 60  
1 MGRSPCEKAHNTKGAATWTEEDRLNIVYHGGCWRSLPKAGLRLCGSCRLRWNY 59  
61 LRPDLKRCNTEEDDELILKSLGKNSLGLRGLRGTRNEIKNWNTRIRKRLIRG 120  
60 LRPDLKRCNTEEDDELILKSLGKNSLGLRGLRGTRNEIKNWNTRIRKRLIRG 119  
121 IDPQTHRLPNTANTNTVTAPELDFRNSPTSVSKSSIKNPSLDFNNEQFKNSTSL 180  
120 IDPQTHRLPNTANTNTVTAPELDFRNSPTSVSKSSIKNPSLDFNNEQFKNSTSL 179  
181 EEPNCASSGNTTDEEOQLHKKOYGPSPNGDINELSGIYSADSSRSNANSASK 240  
180 EEPNCASSGNTTDEEOQLHKKOYGPSPNGDINELSGIYSADSSRSNANSASK 239  
241 PKYDNNPQFLQAOAVAKAVCLCMLGFGTSEICRNCONSNNGFYSCRPIDS 294  
240 PKYDNNPQFLQAOAVAKAVCLCMLGFGTSEICRNCONSNNGFYSCRPIDS 293

RESULT 2

AAG30441  
ID AAG30441 standard; Protein; 257 AA.

AC AAG30441;

DT 17-OCT-2000 (first entry)

DE Arabidopsis thaliana protein fragment SEQ ID NO: 36393.

KM Protein identification; signal transduction pathway; metabolic pathway;  
KW hybridisation assay; genetic mapping; gene expression control; promoter;  
OS termination sequence.

OS Arabidopsis thaliana.

PM EPI033405-A2.

XX

PD 06-SEP-2000.  
PF 25-FEB-2000; 2000EP-0301439.  
XX 25-FEB-1999; 99US-0121825.  
XX 05-MAR-1999; 99US-0123180.  
PR 09-MAR-1999; 99US-0123548.  
PR 23-MAR-1999; 99US-0125788.  
PR 25-MAR-1999; 99US-0126264.  
PR 29-MAR-1999; 99US-0126785.  
PR 01-APR-1999; 99US-0127462.  
PR 06-APR-1999; 99US-0128234.  
PR 08-APR-1999; 99US-0128714.  
PR 16-APR-1999; 99US-0129845.  
PR 19-APR-1999; 99US-0130077.  
PR 21-APR-1999; 99US-0130449.  
PR 23-APR-1999; 99US-0130510.  
PR 23-APR-1999; 99US-0130891.  
PR 28-APR-1999; 99US-0131449.  
PR 30-APR-1999; 99US-0132048.  
PR 30-APR-1999; 99US-0132407.  
PR 04-MAY-1999; 99US-0132484.  
PR 05-MAY-1999; 99US-0132485.  
PR 06-MAY-1999; 99US-0132486.  
PR 06-MAY-1999; 99US-0132487.  
PR 07-MAY-1999; 99US-0132863.  
PR 11-MAY-1999; 99US-0134256.  
PR 14-MAY-1999; 99US-0134218.  
PR 14-MAY-1999; 99US-0134219.  
PR 14-MAY-1999; 99US-0134221.  
PR 14-MAY-1999; 99US-0134370.  
PR 18-MAY-1999; 99US-0134768.  
PR 19-MAY-1999; 99US-0134941.  
PR 20-MAY-1999; 99US-0135124.  
PR 21-MAY-1999; 99US-0135363.  
PR 24-MAY-1999; 99US-0135629.  
PR 25-MAY-1999; 99US-0136021.  
PR 27-MAY-1999; 99US-0136392.  
PR 28-MAY-1999; 99US-0136782.  
PR 01-JUN-1999; 99US-0137232.  
PR 03-JUN-1999; 99US-0137528.  
PR 04-JUN-1999; 99US-0137502.  
PR 07-JUN-1999; 99US-0137724.  
PR 08-JUN-1999; 99US-0138094.  
PR 10-JUN-1999; 99US-0138540.  
PR 10-JUN-1999; 99US-0138847.  
PR 14-JUN-1999; 99US-0139119.  
PR 16-JUN-1999; 99US-0139452.  
PR 16-JUN-1999; 99US-0139453.  
PR 17-JUN-1999; 99US-0139492.  
PR 18-JUN-1999; 99US-0139454.  
PR 18-JUN-1999; 99US-0139455.  
PR 18-JUN-1999; 99US-0139456.  
PR 18-JUN-1999; 99US-0139457.  
PR 18-JUN-1999; 99US-0139458.  
PR 18-JUN-1999; 99US-0139459.  
PR 18-JUN-1999; 99US-0139460.  
PR 18-JUN-1999; 99US-0139461.  
PR 18-JUN-1999; 99US-0139462.  
PR 18-JUN-1999; 99US-0139463.  
PR 18-JUN-1999; 99US-0139750.  
PR 18-JUN-1999; 99US-0139763.  
PR 21-JUN-1999; 99US-0139817.  
PR 22-JUN-1999; 99US-0139899.  
PR 23-JUN-1999; 99US-0140353.  
PR 23-JUN-1999; 99US-0140354.  
PR 24-JUN-1999; 99US-0140354.  
PR 28-JUN-1999; 99US-0140823.  
PR 29-JUN-1999; 99US-0140991.  
PR 30-JUN-1999; 99US-0141287.  
PR 01-JUL-1999; 99US-0141842.  
PR 01-JUL-1999; 99US-0142154.  
PR 02-JUL-1999; 99US-0142055.

PR	23-SEP-1999;	990S-0155486.
PR	24-SEP-1999;	990S-0155659.
PR	28-SEP-1999;	990S-0156458.
PR	29-SEP-1999;	990S-0156596.
PR	04-OCT-1999;	990S-0157117.
PR	05-OCT-1999;	990S-0157753.
PR	06-OCT-1999;	990S-0157865.
PR	07-OCT-1999;	990S-0158029.
PR	08-OCT-1999;	990S-0158232.
PR	12-OCT-1999;	990S-0158369.
PR	13-OCT-1999;	990S-0159293.
PR	13-OCT-1999;	990S-0159294.
PR	13-OCT-1999;	990S-0159295.
PR	14-OCT-1999;	990S-0159329.
PR	14-OCT-1999;	990S-0159330.
PR	14-OCT-1999;	990S-0159331.
PR	14-OCT-1999;	990S-0159637.
PR	14-OCT-1999;	990S-0159638.
PR	18-OCT-1999;	990S-0159584.
PR	21-OCT-1999;	990S-0160741.
PR	21-OCT-1999;	990S-0160767.
PR	21-OCT-1999;	990S-0160768.
PR	21-OCT-1999;	990S-0160770.
PR	21-OCT-1999;	990S-0160814.
PR	21-OCT-1999;	990S-0160815.
PR	22-OCT-1999;	990S-0160980.
PR	22-OCT-1999;	990S-0160981.
PR	22-OCT-1999;	990S-0160989.
PR	25-OCT-1999;	990S-0161404.
PR	25-OCT-1999;	990S-0161405.
PR	25-OCT-1999;	990S-0161406.
PR	26-OCT-1999;	990S-0161359.
PR	26-OCT-1999;	990S-0161360.
PR	26-OCT-1999;	990S-0161361.
PR	28-OCT-1999;	990S-0161920.
PR	28-OCT-1999;	990S-0161992.
PR	28-OCT-1999;	990S-0161993.
PR	29-OCT-1999;	990S-0162142.

Query Match	48.5%	Score 768	DB 21	Length 257
Best Local Similarity	55.6%	Pred. No. 5	9e-65	
Matches 160	Conservative 33	Mismatches 53	Indels 42	Gaps 7
QY	1	MGRSPCEKAHTNKGAWTKREEDQRLINTRYAHGEGCWSLSPKAAGLIRCGSKRLRWNY	60	
Db	1	mgspceekahnmkgawtkreedqllvdylrkngcgwrsllpraagllqrcgskscrllrmny	60	
QY	61	LRPDIKRNFTFEEDBLIIKLSLIGNKWSLLAGLPGRTDNEIRKNWYTHIKRKLISRG	120	
Db	61	lrpdikrnfteeedelliklsllgnkwsllaagrlpyrtdeiknywthkrklisrg	120	
QY	121	IDPQTHRLNQTAMNTVTAPELDFRNSPTSVSSSSSIKNPSLDENYNEFOFSNTDSL	180	
Db	121	idpqnhrline-----svvpspslqndvavetlhdifsgpyrkpepy	160	
QY	181	EEP-----NCTAASGMTTDEEQEQLHKQOYGPSNGODINLELSIG-IVSADSSR-VSN	233	
Db	161	reelgmynvnc-essgtfsek-----ygnsedwvlnlelsvayryestlrkvasv	209	
QY	234	ANSAASKRKVDNNNNQFLQANVAKAVCIQWOLGFTGTSIRICGNQNSN 281		
Db	210	vdsaestrrwgsell---gahesdavclccrlgllfmescrncrvsd 253		
RESULT	3			
AA080529				
ID	AA080529	standard	Protein	236 AA
XX	AA080529			
XX				
XX	17-OCT-2000	(first entry)		
XX				
DE	Arabidopsis thaliana protein fragment SEQ ID NO: 6103			

XX Protein identification: signal transduction pathway; metabolic pathway;  
KM hybridisation assay; genetic mapping; gene expression control; promoter;  
KW termination sequence.  
XX  
OS Arabidopsis thaliana.  
XX  
PN EP1033405-A2.  
XX  
PD 06-SEP-2000.  
XX  
PE 25-FEB-2000; 2000EP-0301439.  
XX  
XX 25-FEB-1999; 99US-0121825.  
PR 05-MAR-1999; 99US-0123140.  
PR 09-MAR-1999; 99US-0123548.  
PR 23-MAR-1999; 99US-0125768.  
PR 25-MAR-1999; 99US-0126264.  
PR 29-MAR-1999; 99US-0126785.  
PR 01-APR-1999; 99US-0127462.  
PR 06-APR-1999; 99US-0128234.  
PR 08-APR-1999; 99US-0128714.  
PR 16-APR-1999; 99US-0129845.  
PR 19-APR-1999; 99US-0130077.  
PR 21-APR-1999; 99US-0130449.  
PR 23-APR-1999; 99US-0130510.  
PR 28-APR-1999; 99US-0130891.  
PR 30-APR-1999; 99US-0131449.  
PR 30-APR-1999; 99US-0132048.  
PR 30-APR-1999; 99US-0132407.  
PR 04-MAY-1999; 99US-0132484.  
PR 05-MAY-1999; 99US-0132485.  
PR 06-MAY-1999; 99US-0132486.  
PR 06-MAY-1999; 99US-0132487.  
PR 07-MAY-1999; 99US-0132863.  
PR 11-MAY-1999; 99US-0134256.  
PR 14-MAY-1999; 99US-0134218.  
PR 14-MAY-1999; 99US-0134219.  
PR 14-MAY-1999; 99US-0134221.  
PR 18-MAY-1999; 99US-0134370.  
PR 18-MAY-1999; 99US-0134768.  
PR 19-MAY-1999; 99US-0134941.  
PR 20-MAY-1999; 99US-0135124.  
PR 21-MAY-1999; 99US-0135353.  
PR 24-MAY-1999; 99US-0135629.  
PR 25-MAY-1999; 99US-0136021.  
PR 27-MAY-1999; 99US-0136382.  
PR 28-MAY-1999; 99US-0136782.  
PR 01-JUN-1999; 99US-0137222.  
PR 03-JUN-1999; 99US-0137528.  
PR 04-JUN-1999; 99US-0137502.  
PR 07-JUN-1999; 99US-0137724.  
PR 08-JUN-1999; 99US-0138094.  
PR 10-JUN-1999; 99US-0138540.  
PR 10-JUN-1999; 99US-0138847.  
PR 14-JUN-1999; 99US-0139119.  
PR 16-JUN-1999; 99US-0139452.  
PR 17-JUN-1999; 99US-0139453.  
PR 17-JUN-1999; 99US-0139452.  
PR 18-JUN-1999; 99US-0139454.  
PR 18-JUN-1999; 99US-0139455.  
PR 18-JUN-1999; 99US-0139456.  
PR 18-JUN-1999; 99US-0139457.  
PR 18-JUN-1999; 99US-0139458.  
PR 18-JUN-1999; 99US-0139459.  
PR 18-JUN-1999; 99US-0139460.  
PR 18-JUN-1999; 99US-0139461.  
PR 18-JUN-1999; 99US-0139462.  
PR 18-JUN-1999; 99US-0139463.  
PR 18-JUN-1999; 99US-0139750.  
PR 18-JUN-1999; 99US-0139763.  
PR 21-JUN-1999; 99US-0139817.  
PR 22-JUN-1999; 99US-0139899.  
  
PR 23-JUN-1999; 99US-0140353.  
PR 23-JUN-1999; 99US-0140354.  
PR 24-JUN-1999; 99US-0140695.  
PR 28-JUN-1999; 99US-0140823.  
PR 29-JUN-1999; 99US-0140991.  
PR 30-JUN-1999; 99US-0141287.  
PR 01-JUL-1999; 99US-0141842.  
PR 01-JUL-1999; 99US-0142154.  
PR 02-JUL-1999; 99US-0142055.  
PR 06-JUL-1999; 99US-0142390.  
PR 08-JUL-1999; 99US-0142803.  
PR 09-JUL-1999; 99US-0142920.  
PR 12-JUL-1999; 99US-0142977.  
PR 13-JUL-1999; 99US-0143542.  
PR 14-JUL-1999; 99US-0143624.  
PR 15-JUL-1999; 99US-0144005.  
PR 16-JUL-1999; 99US-0144085.  
PR 16-JUL-1999; 99US-0144086.  
PR 19-JUL-1999; 99US-0144325.  
PR 19-JUL-1999; 99US-0144331.  
PR 19-JUL-1999; 99US-0144332.  
PR 19-JUL-1999; 99US-0144333.  
PR 19-JUL-1999; 99US-0144334.  
PR 19-JUL-1999; 99US-0144335.  
PR 20-JUL-1999; 99US-0144332.  
PR 20-JUL-1999; 99US-0144632.  
PR 20-JUL-1999; 99US-0144884.  
PR 21-JUL-1999; 99US-0144814.  
PR 21-JUL-1999; 99US-0145086.  
PR 21-JUL-1999; 99US-0145088.  
PR 22-JUL-1999; 99US-0145085.  
PR 22-JUL-1999; 99US-0145087.  
PR 22-JUL-1999; 99US-0145089.  
PR 22-JUL-1999; 99US-0145192.  
PR 23-JUL-1999; 99US-0145145.  
PR 23-JUL-1999; 99US-0145218.  
PR 23-JUL-1999; 99US-0145224.  
PR 26-JUL-1999; 99US-0145276.  
PR 27-JUL-1999; 99US-0145913.  
PR 27-JUL-1999; 99US-0145918.  
PR 27-JUL-1999; 99US-0145919.  
PR 28-JUL-1999; 99US-0145951.  
PR 02-AUG-1999; 99US-0146386.  
PR 02-AUG-1999; 99US-0146388.  
PR 02-AUG-1999; 99US-0146389.  
PR 03-AUG-1999; 99US-0147038.  
PR 04-AUG-1999; 99US-0147204.  
PR 04-AUG-1999; 99US-0147302.  
PR 05-AUG-1999; 99US-0147192.  
PR 05-AUG-1999; 99US-0147260.  
PR 06-AUG-1999; 99US-0147303.  
PR 06-AUG-1999; 99US-0147416.  
PR 09-AUG-1999; 99US-0147493.  
PR 09-AUG-1999; 99US-0147935.  
PR 10-AUG-1999; 99US-0148171.  
PR 11-AUG-1999; 99US-0148319.  
PR 12-AUG-1999; 99US-0148341.  
PR 13-AUG-1999; 99US-0148565.  
PR 13-AUG-1999; 99US-0148684.  
PR 16-AUG-1999; 99US-0149368.  
PR 17-AUG-1999; 99US-0149175.  
PR 18-AUG-1999; 99US-0149426.  
PR 20-AUG-1999; 99US-0149722.  
PR 20-AUG-1999; 99US-0149723.  
PR 20-AUG-1999; 99US-0149929.  
PR 23-AUG-1999; 99US-0149902.  
PR 23-AUG-1999; 99US-0149930.  
PR 25-AUG-1999; 99US-0150566.  
PR 26-AUG-1999; 99US-0150884.  
PR 27-AUG-1999; 99US-0151065.  
PR 27-AUG-1999; 99US-0151066.  
PR 27-AUG-1999; 99US-0151080.  
PR 30-AUG-1999; 99US-0151303.

PR 31-AUG-1999; 99US-0151438.  
 PR 01-SEP-1999; 99US-0151930.  
 PR 07-SEP-1999; 99US-0152366.  
 PR 10-SEP-1999; 99US-0153070.  
 PR 13-SEP-1999; 99US-0153758.  
 PR 15-SEP-1999; 99US-0154018.  
 PR 16-SEP-1999; 99US-0154039.  
 PR 20-SEP-1999; 99US-0154779.  
 PR 22-SEP-1999; 99US-0155139.  
 PR 23-SEP-1999; 99US-0155486.  
 PR 24-SEP-1999; 99US-0155659.  
 PR 28-SEP-1999; 99US-0156458.  
 PR 29-SEP-1999; 99US-0156596.  
 PR 04-OCT-1999; 99US-0157117.  
 PR 05-OCT-1999; 99US-0157753.  
 PR 06-OCT-1999; 99US-0157865.  
 PR 07-OCT-1999; 99US-0158022.  
 PR 08-OCT-1999; 99US-0158232.  
 PR 12-OCT-1999; 99US-0158369.  
 PR 13-OCT-1999; 99US-0158293.  
 PR 13-OCT-1999; 99US-0158294.  
 PR 13-OCT-1999; 99US-0159295.  
 PR 14-OCT-1999; 99US-0159329.  
 PR 14-OCT-1999; 99US-0159330.  
 PR 14-OCT-1999; 99US-0159331.  
 PR 14-OCT-1999; 99US-0159637.  
 PR 14-OCT-1999; 99US-0159638.  
 PR 14-OCT-1999; 99US-0159684.  
 PR 18-OCT-1999; 99US-0159584.  
 PR 21-OCT-1999; 99US-0160747.  
 PR 21-OCT-1999; 99US-0160767.  
 PR 21-OCT-1999; 99US-0160768.  
 PR 21-OCT-1999; 99US-0160770.  
 PR 21-OCT-1999; 99US-0160814.  
 PR 21-OCT-1999; 99US-0160815.  
 PR 22-OCT-1999; 99US-0160980.  
 PR 22-OCT-1999; 99US-0160981.  
 PR 22-OCT-1999; 99US-0160989.  
 PR 25-OCT-1999; 99US-0161404.  
 PR 25-OCT-1999; 99US-0161405.  
 PR 25-OCT-1999; 99US-0161406.  
 PR 26-OCT-1999; 99US-0161359.  
 PR 26-OCT-1999; 99US-0161360.  
 PR 26-OCT-1999; 99US-0161361.  
 PR 28-OCT-1999; 99US-0161920.  
 PR 28-OCT-1999; 99US-0161923.  
 PR 28-OCT-1999; 99US-0161933.  
 PR 29-OCT-1999; 99US-0162142.

Query Match 48.1%; Score 762; DB 21; Length 236;  
 Best Local Similarity 58.9%; Pred. No. 1.9e-64;

Matches 145; Conservative 39; Mismatches 40; Indels 22; Gaps 5;

OY 1 MGRSPCEKAKHNTKNGKMTKEEDORLNIYRVHGGCWRLSPKAGLLRCGRSLRWINY 60  
 DB 1 mgrspceekahntkngkmtkeedordlrytrnhggcwrlspkagllrcgrsrlrwiny 60  
 OY 61 LRPDLKRGNTFEEDLILKLSLGNKWSLIGRLPGRDNEIKNTWTHIKRLISRG 120  
 DB 61 lrpdlkrgntfeeddlilklslgnkwsllgrrlpgrdneiknwtwhikrllshg 120  
 OY 121 IDPQTHRPANQANTWT-VTAPFELDPFRNSPTYSKSSSTIKNPISLDFVYNEFOFSNDS 179  
 DB 121 idpqtthrpqneaklvsqyvvp-----lqndaveysfslavrpkten 163  
 OY 180 LEEPNTASSGNTDEEOQLKHKOOKYGPSNGQ-DINLELSIGIVSDSSRVSNANSAE 238  
 DB 164 sed-ngascsqcttdedlry-ngecyadneghiklniditlfgswsqrlvgysaed 220  
 OY 239 SKPKVD 244  
 DB 221 skpvc 226

RESULT 4  
 AAG49909  
 ID AAG49909 standard; Protein: 236 AA.  
 XX  
 AC AAG49909;  
 XX  
 DT 18-OCT-2000 (first entry)  
 XX  
 DE Arabidopsis thaliana protein fragment SEQ ID NO: 63188.  
 XX  
 KW Protein identification; signal transduction pathway; metabolic pathway;  
 KW hybridisation assay; genetic mapping; gene expression control; promoter;  
 KW termination sequence.  
 XX  
 OS Arabidopsis thaliana.  
 XX  
 PN EP1033405-A2.  
 XX  
 PD 06-SEP-2000.  
 XX  
 PF 25-FEB-2000; 2000EP-0301439.  
 XX  
 PR 25-FEB-1999; 99US-0121825.  
 PR 05-MAR-1999; 99US-0123180.  
 PR 09-MAR-1999; 99US-0123548.  
 PR 23-MAR-1999; 99US-0125788.  
 PR 25-MAR-1999; 99US-0126264.  
 PR 29-MAR-1999; 99US-0126785.  
 PR 01-APR-1999; 99US-0127462.  
 PR 06-APR-1999; 99US-0128234.  
 PR 08-APR-1999; 99US-0128714.  
 PR 16-APR-1999; 99US-0129845.  
 PR 19-APR-1999; 99US-0130077.  
 PR 21-APR-1999; 99US-0130449.  
 PR 23-APR-1999; 99US-0130510.  
 PR 23-APR-1999; 99US-0130891.  
 PR 28-APR-1999; 99US-0131449.  
 PR 30-APR-1999; 99US-0132048.  
 PR 30-APR-1999; 99US-0132407.  
 PR 04-MAY-1999; 99US-0132484.  
 PR 05-MAY-1999; 99US-0132485.  
 PR 06-MAY-1999; 99US-0132486.  
 PR 06-MAY-1999; 99US-0132487.  
 PR 07-MAY-1999; 99US-0132863.  
 PR 11-MAY-1999; 99US-0134256.  
 PR 14-MAY-1999; 99US-0134218.  
 PR 14-MAY-1999; 99US-0134219.  
 PR 14-MAY-1999; 99US-0134221.  
 PR 14-MAY-1999; 99US-0134370.  
 PR 18-MAY-1999; 99US-0134768.  
 PR 19-MAY-1999; 99US-0134941.  
 PR 20-MAY-1999; 99US-0135124.  
 PR 21-MAY-1999; 99US-0135353.  
 PR 24-MAY-1999; 99US-0135629.  
 PR 25-MAY-1999; 99US-0136021.  
 PR 27-MAY-1999; 99US-0136392.  
 PR 28-MAY-1999; 99US-0136782.  
 PR 01-JUN-1999; 99US-0137222.  
 PR 03-JUN-1999; 99US-0137528.  
 PR 04-JUN-1999; 99US-0137502.  
 PR 07-JUN-1999; 99US-0137724.  
 PR 08-JUN-1999; 99US-0138094.  
 PR 10-JUN-1999; 99US-0138540.  
 PR 10-JUN-1999; 99US-0138847.  
 PR 14-JUN-1999; 99US-0139119.  
 PR 16-JUN-1999; 99US-0139452.  
 PR 16-JUN-1999; 99US-0139453.  
 PR 17-JUN-1999; 99US-0139493.  
 PR 18-JUN-1999; 99US-0139454.  
 PR 18-JUN-1999; 99US-0139455.  
 PR 18-JUN-1999; 99US-0139456.  
 PR 18-JUN-1999; 99US-0139457.  
 PR 18-JUN-1999; 99US-0139458.

PR 18-JUN-1999; 99US-0139459.  
PR 18-JUN-1999; 99US-0139460.  
PR 18-JUN-1999; 99US-0139461.  
PR 18-JUN-1999; 99US-0139462.  
PR 18-JUN-1999; 99US-0139463.  
PR 18-JUN-1999; 99US-0139750.  
PR 18-JUN-1999; 99US-0139763.  
PR 21-JUN-1999; 99US-0139817.  
PR 22-JUN-1999; 99US-0139899.  
PR 23-JUN-1999; 99US-0140353.  
PR 23-JUN-1999; 99US-0140354.  
PR 24-JUN-1999; 99US-0140695.  
PR 28-JUN-1999; 99US-0140823.  
PR 29-JUN-1999; 99US-0140991.  
PR 30-JUN-1999; 99US-0141287.  
PR 01-JUL-1999; 99US-0141844.  
PR 01-JUL-1999; 99US-0142154.  
PR 02-JUL-1999; 99US-0142055.  
PR 06-JUL-1999; 99US-0142390.  
PR 08-JUL-1999; 99US-0142803.  
PR 09-JUL-1999; 99US-0142920.  
PR 12-JUL-1999; 99US-0142977.  
PR 13-JUL-1999; 99US-0143542.  
PR 14-JUL-1999; 99US-0143624.  
PR 15-JUL-1999; 99US-0144005.  
PR 16-JUL-1999; 99US-0144085.  
PR 19-JUL-1999; 99US-0144325.  
PR 19-JUL-1999; 99US-0144331.  
PR 19-JUL-1999; 99US-0144332.  
PR 19-JUL-1999; 99US-0144334.  
PR 19-JUL-1999; 99US-0144335.  
PR 20-JUL-1999; 99US-0144352.  
PR 20-JUL-1999; 99US-0144632.  
PR 20-JUL-1999; 99US-0144884.  
PR 21-JUL-1999; 99US-0144814.  
PR 21-JUL-1999; 99US-0145086.  
PR 21-JUL-1999; 99US-0145088.  
PR 22-JUL-1999; 99US-0145085.  
PR 22-JUL-1999; 99US-0145087.  
PR 22-JUL-1999; 99US-0145089.  
PR 22-JUL-1999; 99US-0145192.  
PR 23-JUL-1999; 99US-0145145.  
PR 23-JUL-1999; 99US-0145218.  
PR 23-JUL-1999; 99US-0145224.  
PR 26-JUL-1999; 99US-0145276.  
PR 27-JUL-1999; 99US-0145913.  
PR 27-JUL-1999; 99US-0145918.  
PR 28-JUL-1999; 99US-0145919.  
PR 28-JUL-1999; 99US-0145951.  
PR 02-AUG-1999; 99US-0146386.  
PR 02-AUG-1999; 99US-0146388.  
PR 02-AUG-1999; 99US-0146389.  
PR 03-AUG-1999; 99US-0147038.  
PR 04-AUG-1999; 99US-0147204.  
PR 04-AUG-1999; 99US-0147302.  
PR 05-AUG-1999; 99US-0147192.  
PR 05-AUG-1999; 99US-0147260.  
PR 06-AUG-1999; 99US-0147303.  
PR 06-AUG-1999; 99US-0147416.  
PR 09-AUG-1999; 99US-0147493.  
PR 09-AUG-1999; 99US-0147935.  
PR 10-AUG-1999; 99US-0148171.  
PR 11-AUG-1999; 99US-0148319.  
PR 12-AUG-1999; 99US-0148341.  
PR 13-AUG-1999; 99US-0148365.  
PR 13-AUG-1999; 99US-0148684.  
PR 16-AUG-1999; 99US-0149368.  
PR 17-AUG-1999; 99US-0149175.  
PR 18-AUG-1999; 99US-0149426.  
PR 20-AUG-1999; 99US-0149722.  
PR 20-AUG-1999; 99US-0149723.

PR 20-AUG-1999; 99US-0149929.  
PR 23-AUG-1999; 99US-0149902.  
PR 23-AUG-1999; 99US-0149930.  
PR 25-AUG-1999; 99US-0150566.  
PR 26-AUG-1999; 99US-0150884.  
PR 27-AUG-1999; 99US-0151065.  
PR 27-AUG-1999; 99US-0151066.  
PR 27-AUG-1999; 99US-0151080.  
PR 30-AUG-1999; 99US-0151303.  
PR 31-AUG-1999; 99US-0151438.  
PR 01-SEP-1999; 99US-0151930.  
PR 07-SEP-1999; 99US-0152363.  
PR 10-SEP-1999; 99US-0153070.  
PR 13-SEP-1999; 99US-0153758.  
PR 15-SEP-1999; 99US-0154018.  
PR 16-SEP-1999; 99US-0154039.  
PR 20-SEP-1999; 99US-0154779.  
PR 22-SEP-1999; 99US-0155139.  
PR 23-SEP-1999; 99US-0155886.  
PR 24-SEP-1999; 99US-0155859.  
PR 28-SEP-1999; 99US-0156458.  
PR 29-SEP-1999; 99US-0156596.  
PR 04-OCT-1999; 99US-0157117.  
PR 05-OCT-1999; 99US-0157753.  
PR 06-OCT-1999; 99US-0157865.  
PR 07-OCT-1999; 99US-0158029.  
PR 08-OCT-1999; 99US-0158232.  
PR 12-OCT-1999; 99US-0158369.  
PR 13-OCT-1999; 99US-0159293.  
PR 13-OCT-1999; 99US-0159294.  
PR 13-OCT-1999; 99US-0159295.  
PR 14-OCT-1999; 99US-0159329.  
PR 14-OCT-1999; 99US-0159330.  
PR 14-OCT-1999; 99US-0159331.  
PR 14-OCT-1999; 99US-0159637.  
PR 14-OCT-1999; 99US-0159638.  
PR 18-OCT-1999; 99US-0159584.  
PR 21-OCT-1999; 99US-0160741.  
PR 21-OCT-1999; 99US-0160767.  
PR 21-OCT-1999; 99US-0160768.  
PR 21-OCT-1999; 99US-0160770.  
PR 21-OCT-1999; 99US-0160814.  
PR 21-OCT-1999; 99US-0160815.  
PR 22-OCT-1999; 99US-0160980.  
PR 22-OCT-1999; 99US-0160981.  
PR 22-OCT-1999; 99US-0160989.  
PR 25-OCT-1999; 99US-0161404.  
PR 25-OCT-1999; 99US-0161405.  
PR 25-OCT-1999; 99US-0161406.  
PR 26-OCT-1999; 99US-0161359.  
PR 26-OCT-1999; 99US-0161360.  
PR 26-OCT-1999; 99US-0161361.  
PR 28-OCT-1999; 99US-0161920.  
PR 28-OCT-1999; 99US-0161992.  
PR 28-OCT-1999; 99US-0161993.  
PR 29-OCT-1999; 99US-0162142.

Query Match 48.1%; Score 762; DB 21; Length 236;  
Best Local Similarity 58.9%; Pred. No. 1.9e-64;  
Matches 145; Conservative 39; Mismatches 40; Indels 22; Gaps 5;  
OY 1 MGRSPCEKHAHTNKGMATKEDORLIVIRVHGECNRSLPKAGLLRCGRSCLRWINY 60  
DB 1 mgrspcekahntkngawtkedgrlvdylrnhgscwrsllpcgsacrlrwiny 60  
OY 61 LPDOLKRGNTFEDEDEILILHLSILGNKKSILARLDRGRTNEIKNYNTIHKRLISRG 120  
DB 61 lrpdlkrgnftdedeqlllkhlsilgnkksilagrldprtneiknyntihkrlilshg 120  
OY 121 IDPOTHRPLNOTANTNT-VTAPTELDPRNSPTSVSKSSSIKNPSUDFVYNEFOFSNTDS 179  
DB 121 idpqrhplnotantnt-vtaptelprnsptsvskssiknpsudfvynefofskntds 163

QY 180 LEEPNCTASSGNTDEBOQLHKQOYGPSNGO-DINLELSIGIVSADSSRVSNANSAE 238  
 ID AAB67842 standard; Protein: 236 AA.  
 XX AAB67842;  
 AC AAB67842;  
 XX 29-JUN-2001 (first entry)  
 DE Amino acid sequence of a plant transcription factor G208.  
 XX Transcription factor: flowering time; transgenic plant; vernalisation;  
 KM plant development; plant physiology; flowering.  
 XX Arabidopsis thaliana.  
 OS WO200126459-A2.  
 PN 19-APR-2001.  
 PD 11-OCT-2000; 2000WO-US28141.  
 XX 12-OCT-1999; 99US-0159464.  
 PR 08-NOV-1999; 99US-0164132.  
 PR 17-NOV-1999; 99US-0166228.  
 PR 17-APR-2000; 2000US-0197899.  
 PR 22-AUG-2000; 2000US-0227439.  
 XX (MEND-) MENDEL BIOTECHNOLOGY INC.  
 PA (RATC/) RATCLIFFE O.  
 PA (HEAR/) HEARD J.  
 PA (SAMA/) SAMAHA R.  
 PA (CREE/) CREELMAN R.  
 PA (KEDD/) KEDDIE J.  
 PA (JIANG/) JIANG C.  
 PA (REUB/) REUBER L.  
 PA (RIEC/) RIECHMANN J L.  
 XX Ratcliffe O, Heard J, Samaha R, Creelman R, Keddie J, Jiang C,  
 PI Reuber L, Riechmann JL.  
 XX WPI: 2001-266398/27.  
 DR N-PSDB; AAF80412.  
 XX  
 PT New transgenic plant comprises a recombinant polynucleotide encoding a  
 PT plant transcription factor polypeptide and has a modified flowering  
 PT time or vernalization requirement.  
 XX  
 XX Example 7; Page 105-106; 108pp; English.  
 PS The present sequence represents a plant transcription factor protein  
 CC which modifies the flowering time of a plant. The polynucleotide  
 CC sequence is used to produce transgenic plants which have a modified  
 CC flowering time or a modified vernalisation requirement. The  
 CC polynucleotides and polypeptides are useful for modifying plant  
 CC development, physiology or biochemistry such that the modified plants  
 CC have a trait advantage over wild type plants. In particular they are  
 CC useful for accelerating, delaying or preventing flowering. The  
 CC polynucleotides are also useful as nucleic acid probes and primers.  
 CC They may be used to identify proteins that can modify the activity of  
 CC the transcription factor.  
 XX Sequence 236 AA:

Query Match. 48.1%; Score 762; DB 22; Length 236;  
 Best Local Similarity 58.9%; Pred. No. 1.9e-64;  
 Matches 145; Conservative 39; Mismatches 40; Indels 22; Gaps 5;  
 QY 1 MGRSPCCERAKHTKGAATKEDQRLIVYHVGCGRSLPKAGLRLCGSCRRATNY 60  
 DB 1 mgrspccerakhtkgaatkcedqrlivdyirnhgscwrsipksagllrcgscrlwiny 60  
 QY 61 LRPDLKRGNTDEDELIILKHSILGNKWSLIAGRLPGRFTNEIKNYNTIKRKLIERG 120  
 DB 61 lrpdlkrgntdedeqllilhsilgnkwsliagrlpgrtneiknyntikrkliershg 120  
 QY 121 IDPOTHRPINOATNTNT-VTAPTELDPFRNSPTSVSKSSIKNPISLDFYNEFOKNTDS 179  
 DB 121 idpqtlnrqnesktsesqvvv-----lqndaveyfsfnlavpkten 163  
 QY 180 LEEPNCTASSGNTDEBOQLHKQOYGPSNGO-DINLELSIGIVSADSSRVSNANSAE 238  
 DB 180 lleepnctassgntdeboqlhkqoygpsngo-dinlelsigivsadssrvsnansae 238  
 QY 164 ssd-ngastsgttddedlrg-ngecydsdngshiknldltlfgswsgrlvvgssad 220  
 DB 239 SKPKVD 244  
 DB 221 skpvcdd 226  
 RESULT 6  
 AAB33306  
 ID AAB33306 standard; Protein: 255 AA.  
 XX AAB33306;  
 AC AAB33306;  
 XX 25-JAN-2001 (first entry)  
 DE Pinus radiata transcription factor protein sequence #346.  
 XX  
 XX Pinus radiata transcription factor: gene expression; eucalyptus; pine; acacia;  
 KW poplar; sweetgum; teak; mahogany; b2ip; G-box binding factor;  
 KW basic helix-loop-helix zipper; homeotic; homeodomain; homeobox; MADS;  
 KW homeodomain zipper; LIM domain; AP2; EREBs; zinc finger domain;  
 KW type 2 Cys2His2; CCAAT box element; MYB.  
 XX  
 OS Pinus radiata.  
 PN WO200053724-A2.  
 PD 14-SEP-2000.  
 XX 09-MAR-2000; 2000WO-US06112.  
 PF 11-MAR-1999; 99US-0266513.  
 PR 18-AUG-1999; 99US-0149485.  
 XX  
 PA (GENE-) GENESIS RES & DEV CORP LTD.  
 PA (FLET-) FLETCHER CHALLENGE FORESTS LTD.  
 XX Wood M, McGrath A, Shenk MA, Glenn M,  
 PI Wood M, McGrath A, Shenk MA, Glenn M,  
 XX WPI: 2000-579369/54.  
 DR  
 XX  
 PT New isolated polynucleotide encoding a plant transcription factor for  
 PT producing a plant e.g. a woody plant, preferably eucalyptus or pine,  
 PT having modified gene expression or modified activity of a polypeptide  
 PT  
 XX  
 XX Claim 8; Pages 717-718; 747pp; English.  
 PS The present invention relates to novel plant transcription factors from  
 CC Eucalyptus grandis or Pinus radiata. The present sequence is one such  
 CC transcription factor. The transcription factor may be used to produce a  
 CC plant having modified gene expression such as a woody plant e.g. a  
 CC eucalyptus, pine, acacia, poplar, sweetgum, teak, or mahogany species or  
 CC to modify the activity of a polypeptide in a plant. The transcription  
 CC factors of the present invention are members from the following families

CC of regulatory proteins: bzip, bzip family of G-box binding factors, basic  
 CC helix-loop-helix zipper, homeodomain/homeobox/MADS, homeodomain  
 CC zipper, LIM domain, AP2 and EREBS, zinc finger domains of type 2  
 CC Cys2His2, COAT box elements and MYB.

XX Sequence 255 AA;

Query Match 46.0%; Score 728.5; DB 21; Length 255;  
 Best Local Similarity 51.7%; Pred. No. 3,4e-61;  
 Matches 155; Conservative 32; Mismatches 52; Indels 61; Gaps 10;

QY 1 MGRSPCEKATNTNGAMTKREDOIRINIRVHGECMSLPRAGALLRGKSCRLWIMV 60  
 |||||  
 Db 1 mgtrspcekatntngawtkreedrlihlrttgygcwslpkaaglmrcgscrltwiny 60  
 |||||  
 QY 61 LRPLKRGNFEEDELLKLSLGNKMSLAGRLPGRDNEIKRYWNTIKRLISRG 120  
 |||||  
 Db 61 lrpldkrgnfseedeelviklshlgnkwsliagrlpgrtdneikrywntihkrliny 120  
 |||||  
 QY 121 IDPQHRPLNQTANTNTVTAPT-----ELDFRNSPTSVSKSSIKNPGL-DFNYNEQF 173  
 |||||  
 Db 121 ldpqshrpplgphnsnt-tcpslpalehellvfgfp-----rtpeladtf---fgy 166  
 |||||  
 QY 174 KSNMDSLEPRCTASSGWTDEEOQLHKKQYGPSPNGODINLEISGIVSADSRVSN 233  
 : : : : :  
 Db 167 ertesspmep-----atckdaeeh-----pdlnldiclsipvhsppatsr 206  
 : : : : :  
 QY 234 ANS-----AESKPKYDNNNFOLFQAMVAKAVCLCQOLGFTSEICNCONSNS--NGEYVS 287  
 : : : : :  
 Db 207 aswvdgcvdakpn-----svschmglyqnygvqcentyccesaasgvsfyt 252  
 : : : : :  
 RESULT 7  
 AAC32956  
 ID AAC32956 standard; Protein: 282 AA.

AC AAC32956;  
 DT 17-OCT-2000 (first entry)

DE Arabidopsis thaliana protein fragment SEQ ID NO: 39857.

KW Protein identification; signal transduction pathway; metabolic pathway;  
 KW hybridisation assay; genetic mapping; gene expression control; promoter;  
 KW termination sequence.

OS Arabidopsis thaliana.

XX  
 XX EP1033405-A2.

PD 06-SEP-2000.

PF 25-FEB-2000; 2000EP-0301439.

XX 25-FEB-1999; 99US-0121825.  
 XX 05-MAR-1999; 99US-0123180.  
 XX 09-MAR-1999; 99US-0123548.  
 XX 23-MAR-1999; 99US-0125788.  
 XX 25-MAR-1999; 99US-0126264.  
 XX 29-MAR-1999; 99US-0126785.  
 XX 01-APR-1999; 99US-0127462.  
 XX 06-APR-1999; 99US-0128234.  
 XX 08-APR-1999; 99US-0128714.  
 XX 16-APR-1999; 99US-0129845.  
 XX 19-APR-1999; 99US-0130077.  
 XX 21-APR-1999; 99US-0130449.  
 XX 23-APR-1999; 99US-0130510.  
 XX 23-APR-1999; 99US-0130891.  
 XX 28-APR-1999; 99US-0131449.  
 XX 30-APR-1999; 99US-0132048.  
 XX 30-APR-1999; 99US-0132407.  
 XX 04-MAY-1999; 99US-0132484.

PR 05-MAY-1999; 99US-0132485.  
 PR 06-MAY-1999; 99US-0132486.  
 PR 06-MAY-1999; 99US-0132487.  
 PR 07-MAY-1999; 99US-0132863.  
 PR 11-MAY-1999; 99US-0134256.  
 PR 14-MAY-1999; 99US-0134218.  
 PR 14-MAY-1999; 99US-0134219.  
 PR 14-MAY-1999; 99US-0134221.  
 PR 14-MAY-1999; 99US-0134370.  
 PR 18-MAY-1999; 99US-0134768.  
 PR 19-MAY-1999; 99US-0134941.  
 PR 20-MAY-1999; 99US-0135124.  
 PR 21-MAY-1999; 99US-0135353.  
 PR 24-MAY-1999; 99US-0135629.  
 PR 25-MAY-1999; 99US-0136021.  
 PR 27-MAY-1999; 99US-0136392.  
 PR 28-MAY-1999; 99US-0136782.  
 PR 01-JUN-1999; 99US-0137222.  
 PR 03-JUN-1999; 99US-0137528.  
 PR 04-JUN-1999; 99US-0137502.  
 PR 07-JUN-1999; 99US-0137724.  
 PR 08-JUN-1999; 99US-0138094.  
 PR 10-JUN-1999; 99US-0138540.  
 PR 10-JUN-1999; 99US-0138847.  
 PR 14-JUN-1999; 99US-0139119.  
 PR 16-JUN-1999; 99US-0139452.  
 PR 16-JUN-1999; 99US-0139453.  
 PR 17-JUN-1999; 99US-0139492.  
 PR 18-JUN-1999; 99US-0139454.  
 PR 18-JUN-1999; 99US-0139455.  
 PR 18-JUN-1999; 99US-0139455.  
 PR 18-JUN-1999; 99US-0139457.  
 PR 18-JUN-1999; 99US-0139458.  
 PR 18-JUN-1999; 99US-0139459.  
 PR 18-JUN-1999; 99US-0139460.  
 PR 18-JUN-1999; 99US-0139461.  
 PR 18-JUN-1999; 99US-0139462.  
 PR 18-JUN-1999; 99US-0139463.  
 PR 18-JUN-1999; 99US-0139750.  
 PR 18-JUN-1999; 99US-0139763.  
 PR 21-JUN-1999; 99US-0139817.  
 PR 22-JUN-1999; 99US-0139889.  
 PR 23-JUN-1999; 99US-0140353.  
 PR 23-JUN-1999; 99US-0140354.  
 PR 24-JUN-1999; 99US-0140695.  
 PR 28-JUN-1999; 99US-0140823.  
 PR 29-JUN-1999; 99US-0140991.  
 PR 30-JUN-1999; 99US-0141287.  
 PR 01-JUL-1999; 99US-0141842.  
 PR 01-JUL-1999; 99US-0142154.  
 PR 02-JUL-1999; 99US-0142055.  
 PR 06-JUL-1999; 99US-0142390.  
 PR 08-JUL-1999; 99US-0142803.  
 PR 09-JUL-1999; 99US-0142920.  
 PR 12-JUL-1999; 99US-0142977.  
 PR 13-JUL-1999; 99US-0143542.  
 PR 14-JUL-1999; 99US-0143624.  
 PR 15-JUL-1999; 99US-0144005.  
 PR 16-JUL-1999; 99US-0144085.  
 PR 16-JUL-1999; 99US-0144086.  
 PR 19-JUL-1999; 99US-0144325.  
 PR 19-JUL-1999; 99US-0144331.  
 PR 19-JUL-1999; 99US-0144332.  
 PR 19-JUL-1999; 99US-0144333.  
 PR 19-JUL-1999; 99US-0144334.  
 PR 19-JUL-1999; 99US-0144334.  
 PR 19-JUL-1999; 99US-0144335.  
 PR 20-JUL-1999; 99US-0144352.  
 PR 20-JUL-1999; 99US-0144632.  
 PR 20-JUL-1999; 99US-0144884.  
 PR 21-JUL-1999; 99US-0144814.  
 PR 21-JUL-1999; 99US-0145086.  
 PR 21-JUL-1999; 99US-0145088.  
 PR 22-JUL-1999; 99US-0145085.

```

PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 28-JUL-1999; 99US-0145919.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149923.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.

```

```

PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 22-OCT-1999; 99US-0160989.
PR 22-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 26-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 45.68; Score 723; DB 21; Length 282;
Best Local Similarity 52.08; Pred. No. 1.3e-60;
Matches 146; Conservative 38; Mismatches 49; Indels 48; Gaps 7;

QY 1 MGRSPCEKATNKGAWTKEEDRLINIRVHGECCWRSIPRAAGLLRCGKSCRLRWNY 60
DB 1 mgsrpscekatnkgawtkederlvayikangecwrsipraagllrcgkscrlrwny 60
QY 61 LRPDLKRGNTFEDEDELLIKHSLGKNKWSLIAGRLPGFTDNEIKNYWTHIRKILSRG 120
DB 61 lrpdlkrgntfeedeellikhsllgnkwsliagrtpgtndelnknywthirkilnrg 120
QY 121 IDPQTHPLNCTANTNTYTAPLEDFRNSPT-SVSKSSSIRKPSLDFNNEPFRKNTDS 179
DB 121 idpqrhplnctantntytapledfnrnspt-svsksssirkpssldfnnepfrkntds 179
QY 121 idprshrp1-qesassqskptqlpevrsntlnisftsa--pkvelfhesisfpqskse 176
DB 121 idprshrp1-qesassqskptqlpevrsntlnisftsa--pkvelfhesisfpqskse 176
QY 180 LEEPNCRTASSGMTNDEEOEOELHKROXGSPNSGODINLEL----- 219
DB 180 leepncrtassgmtndeoeoeelhkroxxgspnsngodinlel----- 219
QY 177 1-----smllfkeekdeepvgekt-----pdlnlelrslpddvdrilqghkstcp 222
DB 177 1-----smllfkeekdeepvgekt-----pdlnlelrslpddvdrilqghkstcp 222
QY 220 -----SIGIVSADSSRVSNAN-----SASESKPKVDNNNFOPL 251
DB 220 -----sigivsadssrvsnan-----sasekpkvddnnnfopl 251
QY 223 rcfkcsiglmngmeccrgmrcdvvgssaksgdmsngftfl 263
DB 223 rcfkcsiglmngmeccrgmrcdvvgssaksgdmsngftfl 263

RESULT 8
AAE02476
ID AAE02476 standard; Proteio; 282 AA.
XX
AC AAE02476;
XX
DE 10-AUG-2001 (first entry)
XX
DE Arabidopsis thaliana transcription factor G664.
XX
KW Transcription factor; environmental stress tolerance; gene therapy;
KW plant structure; plant development.
XX
OS Arabidopsis thaliana.
XX
FH Key
FT Domain Location/Qualifiers
FT Domain 13..116
FT /note="Conserved domain"
XX
PN W0200136598-A1.
XX
PF 14-NOV-2000; 2000WO-US31458.
XX
PD 25-MAY-2001.
XX
PE 14-NOV-2000; 2000WO-US31458.
XX
PR 17-NOV-1999; 99US-0166228.
PR 17-APR-2000; 2000US-0197899.
PR 22-AUG-2000; 2000US-0227439.
XX
PA (MEND-) MENDEL BIOTECHNOLOGY INC.
PA (PINE/) PINEDA O.

```



PA (YUGG/) YU G.  
 PA (CREE/) CREELMAN R.  
 PA (RIEC/) RIECHMANN J L.  
 PA (HEAR/) HEARD J.  
 PA (RATC/) RATCLIFFE O.  
 PA (REUB/) REUBER L.  
 PA (KEDD/) KEDDIE J.  
 XX  
 PI Pineda O, Yu G, Creelman R, Riechmann JL, Heard J, Ratcliffe O,  
 PI Reuber L, Keddie J;  
 XX  
 DR MPI; 2001-336000/35.  
 DR N-PSDB; AAD06456.  
 XX  
 PT Nucleic acids encoding plant transcription factor polypeptides, useful  
 PT for altering the environmental stress tolerance characteristics of  
 PT plants -  
 XX  
 PS Claim 4; Page 83-84; 116pp; English.  
 XX  
 CC The present amino acid sequence is Arabidopsis thaliana transcription  
 CC factor. This novel transcription factor is useful for  
 CC modifying a plant's phenotype in desirable ways, such as modifying a  
 CC plants environmental stress. The transcription factor is encoded by  
 CC environmental stress tolerance gene derived from Arabidopsis thaliana.  
 CC The transcription factors and the genes encoding them are used to alter  
 CC the structure and developmental characteristics of plants such as  
 CC soybean, wheat, corn, potato, cotton, rice, oilseed rape, sunflower,  
 CC alfalfa, sugarcane, turf, banana, blackberry, blueberry, strawberry,  
 CC raspberry, cantaloupe, carrot, cauliflower, coffee, onion, cucumber,  
 CC eggplant, grapes, honey dew, lettuce, mango, melon, papaya, peas,  
 CC peppers, pineapple, spinach, squash, sweet corn, tobacco, tomato,  
 CC watermelon, roseaceous fruits and/or vegetable brassicas. These sequences  
 CC are also used for modifying traits associated with environmental stress  
 CC tolerance, such as freezing, chilling, heat, drought, water saturation,  
 CC salt, photoconditions, radiation and ozone. The transcription factors  
 CC are used in gene therapy.  
 CC  
 XX  
 XX Sequence 282 AA;  
 SQ  
 Query Match 45.68; Score 723; DB 22; Length 282;  
 Best Local Similarity 52.08; Pred. No. 1.3e-60;  
 Matches 146; Conservative 38; Mismatches 49; Indels 48; Gaps 7;  
 QY 1 MGRSPCEKATNKGAWKKEEDRLINIRVHGECWRLPRKAGLLRCGKSCRLRWINY 60  
 Db 1 mgrspcekatnkgawkkee derlvayjlkahgecwrlpkaagllrcgkscrlrwiny 60  
 QY 61 LRPLKRGNTFEDEDELIKLSLGNKWSLIAGRLPGRTDNEIKNYWNTIRKRLISRG 120  
 Db 61 lrplkrgntfeede delikls slgnkwsliagrlpg rtdneiknywntir krlisrg 120  
 QY 121 IDPOTHRPLNTANTNTATTELDFRNSPT-SVSKSSSIKPSLDPNNEFPKSTDS 179  
 Db 121 idpshrp lntantntat teldfrnspt-svskss si kpsldp nnefpk stds 179  
 QY 180 LEEFNCTAASGTTDEEOQLHKKQYGPNSGODINLE----- 219  
 Db 177 l-----sm ltfk eekdec pygek f-----pdlnle lrlslp dddvdr lqghyksttp 222  
 QY 220 -----SIGIYASDSSRYSNAN---SASCKPKVVDNNNFQFL 251  
 Db 223 rcfksclgmlngmeo rcmrcdvvgsgsksgdmangfdfl 263  
 RESULT 9  
 AAG30442  
 ID AAG30442 standard; Protein: 246 AA.  
 XX  
 AC AAG30442;  
 XX  
 DT 17-OCT-2000 (first entry)

XX  
 DE Arabidopsis thaliana protein fragment. SEQ ID NO: 36394.  
 XX  
 KW Protein identification; signal transduction pathway; metabolic pathway;  
 KW hybridisation assay; genetic mapping; gene expression control; promoter;  
 KW termination sequence.  
 XX  
 OS Arabidopsis thaliana.  
 XX  
 PN EP1033405-A2.  
 PN  
 PD 06-SEP-2000.  
 XX  
 XX 25-FEB-2000; 2000EP-0301439.  
 XX  
 PR 25-FEB-1999; 99US-0121825.  
 PR 05-MAR-1999; 99US-0123180.  
 PR 09-MAR-1999; 99US-0123548.  
 PR 23-MAR-1999; 99US-0125788.  
 PR 25-MAR-1999; 99US-0126264.  
 PR 29-MAR-1999; 99US-0126785.  
 PR 01-APR-1999; 99US-0127462.  
 PR 06-APR-1999; 99US-0128234.  
 PR 08-APR-1999; 99US-0128714.  
 PR 16-APR-1999; 99US-0129845.  
 PR 19-APR-1999; 99US-0130077.  
 PR 21-APR-1999; 99US-0130449.  
 PR 23-APR-1999; 99US-0130510.  
 PR 28-APR-1999; 99US-0130891.  
 PR 30-APR-1999; 99US-0131449.  
 PR 30-APR-1999; 99US-0132048.  
 PR 30-APR-1999; 99US-0132407.  
 PR 04-MAY-1999; 99US-0132484.  
 PR 05-MAY-1999; 99US-0132485.  
 PR 06-MAY-1999; 99US-0132486.  
 PR 06-MAY-1999; 99US-0132487.  
 PR 07-MAY-1999; 99US-0132863.  
 PR 11-MAY-1999; 99US-0134256.  
 PR 14-MAY-1999; 99US-0134218.  
 PR 14-MAY-1999; 99US-0134219.  
 PR 14-MAY-1999; 99US-0134221.  
 PR 14-MAY-1999; 99US-0134370.  
 PR 18-MAY-1999; 99US-0134768.  
 PR 19-MAY-1999; 99US-0134941.  
 PR 20-MAY-1999; 99US-0135124.  
 PR 21-MAY-1999; 99US-0135353.  
 PR 21-MAY-1999; 99US-0135629.  
 PR 25-MAY-1999; 99US-0136021.  
 PR 27-MAY-1999; 99US-0136392.  
 PR 28-MAY-1999; 99US-0136782.  
 PR 01-JUN-1999; 99US-0137222.  
 PR 03-JUN-1999; 99US-0137528.  
 PR 04-JUN-1999; 99US-0137502.  
 PR 07-JUN-1999; 99US-0137724.  
 PR 08-JUN-1999; 99US-0138094.  
 PR 10-JUN-1999; 99US-0138540.  
 PR 10-JUN-1999; 99US-0138847.  
 PR 14-JUN-1999; 99US-0139119.  
 PR 16-JUN-1999; 99US-0139452.  
 PR 16-JUN-1999; 99US-0139453.  
 PR 17-JUN-1999; 99US-0139454.  
 PR 18-JUN-1999; 99US-0139455.  
 PR 18-JUN-1999; 99US-0139456.  
 PR 18-JUN-1999; 99US-0139457.  
 PR 18-JUN-1999; 99US-0139458.  
 PR 18-JUN-1999; 99US-0139459.  
 PR 18-JUN-1999; 99US-0139460.  
 PR 18-JUN-1999; 99US-0139461.  
 PR 18-JUN-1999; 99US-0139462.  
 PR 18-JUN-1999; 99US-0139463.  
 PR 18-JUN-1999; 99US-0139750.  
 PR 18-JUN-1999; 99US-0139763.

```

PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139859.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 02-JUL-1999; 99US-0142154.
PR 06-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.

```

```

PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0158655.
PR 07-OCT-1999; 99US-0158829.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

```

## Query Match

44.3%; Score 702; DB 21; Length 246;  
 Best Local Similarity 54.0%; Pred. No. 1, le-58;  
 Matches 149; Conservative 33; Mismatches 52; Indels 42; Gaps 7;

```

QY 13 NKGAMTKEDORLINIRYHGEGRSLPRAAGILRGKSGRLWVINYLRDILKGNFTE 72
DB 2 nkgawtkedqllvdylrkrggcwslpraagqrgcgsrlwmylrrpdlrkgnfte 61
QY 73 EDELLIKLHSLGNKWSLIRGRLPGRTDNEIKYVWTHIRKRLISRGIDQTHRPLNOT 132
DB 62 eedelliklshslgnkwsllrgrlptdneikyvthlrrklrlsrgidqthrplnot 120
QY 133 ANTNVTAFTLEDFRNSPVSXSSSTIKNPSLDENYNEFOFKSMTDSLEEP-----NCTA 187
DB 121 -----svpspslqndvvelthldfsgpvkpepyreelgmwnc-e 160
QY 188 SSGMTDEEOEOELARKKOQYPSNGODINLELSIG-IVSADSSR-VSNVANSAESKPKVDN 245
DB 161 ssgtsek-----ygnedwvlnlelsvqpyryestktkvavdsaestrwgs 210
QY 246 NNFQFLQAMVAKAVCLCWLQGLGFTSICRQNSN 281
DB 211 elf---ghesdavlccrlgflfnesrcnrcvsvd 242

```

	RESULT	10
AAB33322	ID	AAB33322 standard; Protein: 225 AA.
XX	AC	
XX	AA	AAB33322;
XX	D7	25-JAN-2001 (first entry)
DE		
XX		Pinus radiata transcription factor protein sequence #361.
KW		Plant; transcription factor; gene expression; eucalyptus; pine; acacia;
KM		poplar; sweetgum; teak; mahogany; bZIP; G-box binding factor;
KV		basic helix-loop-helix zipper; homeotic; homeodomain; homeobox; MADS;
RK		homeodomain zipper; LIM domain; AP2; ERBS; zinc finger domain;
-RK		Type 2 Cys2His2; CCAAT box element; MYB.
OS		
SS		Pinus radiata.
XN		
PN		MO200053724-A2.
PD		
XX		14-SEP-2000.
PF		09-MAR-2000; 2000WO-US06112.
XX		
PR		11-MAR-1999; 99US-0266513.
PR		18-AUG-1999; 99US-0149485.
XX		
PA		(GENE-) GENESIS RES & DEV CORP LTD.
PL		(FLET-) FLETCHEER CHALLENGE FORESTS LTD.
XX		
PI		Wood M, McGrath A, Shenk MA, Glenn M;
DR		WPJ; 2000-579369/54.
XX		
PT		New isolated polynucleotide encoding a plant transcription factor for
PT		producing a plant e.g. a woody plant, preferably eucalyptus or pine,
FT		having modified gene expression or modified activity of a polypeptide
PS		.
XX		Claim 8; Pages 724-725; 747pp; English.
CC		The present invention relates to novel plant transcription factors from
CC		Eucalyptus grandis or Pinus radiata. The present sequence is one such
CC		a plant having modified gene expression such as a woody plant e.g. a
CC		eucalyptus, pine, acacia, poplar, sweetgum, teak, or mahogany species or
CC		factors of the present invention are members of the following families
CC		of regulatory proteins: bZIP, bZIP family of G-box binding factors, basic
CC		helix-loop-helix zipper, homeotic/homeodomain/homeobox/MADS, homeodomain
CC		zipper, LIM domain, AP2 and EREBs, zinc finger domains of type 2
CC		Cys2His2, CCAAT box elements and MYB.
SQ		Sequence 225 AA;
OY		Query Match 44.3%; Score 701; DB 21; Length 225;
		Best Local Similarity 54.5%; Pred.No. 1.2e-58;
MATCHES	146;	Conservative 26; Mismatches 48; Indels 48; Gaps 7;
OY	1	MGRSCECAHTNKAMFKKEEDORILNTIRVGEQMSLPPAAGLLRCGSKRLWTNY 60
DB	1	mgrsccecahtnkgawtkgeddrilantirahgegwslpkaaillrcgskrlwtiny 60
OY	61	LRPDIKRGNFTEEBDELIIKLHSLIGNKWSLIAGRLPRTDNEIKNNWNTHIKRLLISRG 120
DB	61	lrpdikrgsfteebdeliiiklhfsgnwksllagrlprtdeinlnwnthikrlliskg 120
OY	121	IDPCRHRLNLCTANNVTAPF---ELDFRNSPITSVSASSISKNSLDENRYEPQFSMT 177
DB	121	ldpcrhrlngpnnt-pvtrpvleheipafgnpatpeladllqhnrl----- 167

Oy	178 DSELEFNCASSGMTDDEEOEOLHKKQYCFPSNGODINLELSIGIVASDRYSVANSAA	237
Dd	168 ---ssplkpaas---daeh-----pdlnlhcslspnsapavrvssv	206
Oy	238 ESKPKVDNNNFOLFQAMVAKAVCLCWO	265
Dd	207 dt--tydsms-----nsgdlcwg	223
	RESULT 11	
ID	AAG31297	
XX	AAG31297 standard; Protein; 274 AA.	
XX		
AC	AAG31297;	
DT	17-OCT-2000 (first entry)	
XX		
DE	Arabidopsis thaliana protein fragment SEQ ID NO: 37561.	
KW	Protein identification; signal transduction pathway; metabolic pathway;	
KM	hybridisation assay; genetic mapping; gene expression control; promoter	
XX	termination sequence.	
OS	Arabidopsis thaliana.	
PN	EP1033405-A2.	
XX		
PD	06-SEP-2000.	
PF	25-FEB-2000; 2000EP-0301439.	
XX		
PR	25-FEB-1999; 99US-0121825.	
PR	05-MAR-1999; 99US-0123180.	
PR	09-MAR-1999; 99US-0123548.	
PR	23-MAR-1999; 99US-0125788.	
PR	25-MAR-1999; 99US-0126264.	
PR	29-MAR-1999; 99US-0126785.	
PR	01-APR-1999; 99US-0127462.	
PR	06-APR-1999; 99US-0128234.	
PR	08-APR-1999; 99US-0128714.	
PR	16-APR-1999; 99US-0129845.	
PR	19-APR-1999; 99US-0130077.	
PR	21-APR-1999; 99US-0130449.	
PR	23-APR-1999; 99US-0130510.	
PR	28-APR-1999; 99US-0130891.	
PR	30-APR-1999; 99US-0131449.	
PR	30-APR-1999; 99US-0132048.	
PR	04-MAY-1999; 99US-0132407.	
PR	05-MAY-1999; 99US-0132484.	
PR	06-MAY-1999; 99US-0132485.	
PR	06-MAY-1999; 99US-0132486.	
PR	06-MAY-1999; 99US-0132487.	
PR	07-MAY-1999; 99US-0132863.	
PR	11-MAY-1999; 99US-0134256.	
PR	14-MAY-1999; 99US-0134218.	
PR	14-MAY-1999; 99US-0134219.	
PR	14-MAY-1999; 99US-0134221.	
PR	14-MAY-1999; 99US-0134370.	
PR	18-MAY-1999; 99US-0134768.	
PR	19-MAY-1999; 99US-0134941.	
PR	20-MAY-1999; 99US-0135124.	
PR	21-MAY-1999; 99US-0135353.	
PR	24-MAY-1999; 99US-0135629.	
PR	25-MAY-1999; 99US-0136021.	
PR	27-MAY-1999; 99US-0136392.	
PR	28-MAY-1999; 99US-0136782.	
PR	01-JUN-1999; 99US-0137222.	
PR	03-JUN-1999; 99US-0137528.	
PR	04-JUN-1999; 99US-0137582.	
PR	07-JUN-1999; 99US-0137724.	
PR	08-JUN-1999; 99US-0138094.	
PR	10-JUN-1999; 99US-0138540.	

PR 10-JUN-1999; 99US-0138847.  
PR 14-JUN-1999; 99US-0139119.  
PR 16-JUN-1999; 99US-0139452.  
PR 17-JUN-1999; 99US-0139453.  
PR 18-JUN-1999; 99US-0139454.  
PR 18-JUN-1999; 99US-0139455.  
PR 18-JUN-1999; 99US-0139456.  
PR 18-JUN-1999; 99US-0139457.  
PR 18-JUN-1999; 99US-0139458.  
PR 18-JUN-1999; 99US-0139459.  
PR 18-JUN-1999; 99US-0139460.  
PR 18-JUN-1999; 99US-0139461.  
PR 18-JUN-1999; 99US-0139462.  
PR 18-JUN-1999; 99US-0139463.  
PR 18-JUN-1999; 99US-0139750.  
PR 18-JUN-1999; 99US-0139763.  
PR 21-JUN-1999; 99US-0139817.  
PR 22-JUN-1999; 99US-0139889.  
PR 23-JUN-1999; 99US-0140353.  
PR 23-JUN-1999; 99US-0140354.  
PR 24-JUN-1999; 99US-0140695.  
PR 28-JUN-1999; 99US-0140823.  
PR 29-JUN-1999; 99US-0140981.  
PR 30-JUN-1999; 99US-0141287.  
PR 01-JUL-1999; 99US-0141842.  
PR 02-JUL-1999; 99US-0142154.  
PR 02-JUL-1999; 99US-0142055.  
PR 06-JUL-1999; 99US-0142390.  
PR 08-JUL-1999; 99US-0142803.  
PR 09-JUL-1999; 99US-0142920.  
PR 12-JUL-1999; 99US-0142977.  
PR 13-JUL-1999; 99US-0143542.  
PR 14-JUL-1999; 99US-0143624.  
PR 15-JUL-1999; 99US-0144005.  
PR 16-JUL-1999; 99US-0144085.  
PR 16-JUL-1999; 99US-0144086.  
PR 19-JUL-1999; 99US-0144325.  
PR 19-JUL-1999; 99US-0144331.  
PR 19-JUL-1999; 99US-0144332.  
PR 19-JUL-1999; 99US-0144333.  
PR 19-JUL-1999; 99US-0144334.  
PR 19-JUL-1999; 99US-0144335.  
PR 20-JUL-1999; 99US-0144352.  
PR 20-JUL-1999; 99US-0144632.  
PR 20-JUL-1999; 99US-0144884.  
PR 21-JUL-1999; 99US-0144884.  
PR 21-JUL-1999; 99US-0145086.  
PR 21-JUL-1999; 99US-0145088.  
PR 22-JUL-1999; 99US-0145085.  
PR 22-JUL-1999; 99US-0145087.  
PR 22-JUL-1999; 99US-0145089.  
PR 22-JUL-1999; 99US-0145192.  
PR 23-JUL-1999; 99US-0145145.  
PR 23-JUL-1999; 99US-0145218.  
PR 23-JUL-1999; 99US-0145224.  
PR 26-JUL-1999; 99US-0145276.  
PR 27-JUL-1999; 99US-0145913.  
PR 27-JUL-1999; 99US-0145918.  
PR 27-JUL-1999; 99US-0145919.  
PR 28-JUL-1999; 99US-0146386.  
PR 02-AUG-1999; 99US-0146386.  
PR 02-AUG-1999; 99US-0146388.  
PR 03-AUG-1999; 99US-0146389.  
PR 03-AUG-1999; 99US-0147038.  
PR 04-AUG-1999; 99US-0147204.  
PR 04-AUG-1999; 99US-0147302.  
PR 05-AUG-1999; 99US-0147192.  
PR 05-AUG-1999; 99US-0147260.  
PR 06-AUG-1999; 99US-0147303.  
PR 06-AUG-1999; 99US-0147416.  
PR 09-AUG-1999; 99US-0147493.  
PR 09-AUG-1999; 99US-0147935.

PR 10-AUG-1999; 99US-0148171.  
PR 11-AUG-1999; 99US-0148319.  
PR 12-AUG-1999; 99US-0148341.  
PR 13-AUG-1999; 99US-0148565.  
PR 13-AUG-1999; 99US-0148684.  
PR 16-AUG-1999; 99US-0148368.  
PR 17-AUG-1999; 99US-0148175.  
PR 18-AUG-1999; 99US-0148426.  
PR 20-AUG-1999; 99US-0149722.  
PR 20-AUG-1999; 99US-0149723.  
PR 20-AUG-1999; 99US-0149929.  
PR 23-AUG-1999; 99US-0149902.  
PR 23-AUG-1999; 99US-0149930.  
PR 25-AUG-1999; 99US-0150566.  
PR 26-AUG-1999; 99US-0150884.  
PR 27-AUG-1999; 99US-0151065.  
PR 27-AUG-1999; 99US-0151066.  
PR 27-AUG-1999; 99US-0151080.  
PR 30-AUG-1999; 99US-0151303.  
PR 31-AUG-1999; 99US-0151438.  
PR 01-SEP-1999; 99US-0151930.  
PR 07-SEP-1999; 99US-0152363.  
PR 10-SEP-1999; 99US-0153070.  
PR 13-SEP-1999; 99US-0153758.  
PR 15-SEP-1999; 99US-0154018.  
PR 16-SEP-1999; 99US-0154039.  
PR 20-SEP-1999; 99US-0154779.  
PR 22-SEP-1999; 99US-0155139.  
PR 23-SEP-1999; 99US-0155486.  
PR 24-SEP-1999; 99US-0155659.  
PR 28-SEP-1999; 99US-0156458.  
PR 29-SEP-1999; 99US-0156596.  
PR 04-OCT-1999; 99US-0157117.  
PR 05-OCT-1999; 99US-0157753.  
PR 06-OCT-1999; 99US-0157865.  
PR 07-OCT-1999; 99US-0158029.  
PR 08-OCT-1999; 99US-0158322.  
PR 12-OCT-1999; 99US-0158369.  
PR 13-OCT-1999; 99US-0159293.  
PR 13-OCT-1999; 99US-0159294.  
PR 13-OCT-1999; 99US-0159295.  
PR 14-OCT-1999; 99US-0159329.  
PR 14-OCT-1999; 99US-0159330.  
PR 14-OCT-1999; 99US-0159331.  
PR 14-OCT-1999; 99US-0159637.  
PR 14-OCT-1999; 99US-0159638.  
PR 18-OCT-1999; 99US-0159584.  
PR 21-OCT-1999; 99US-0160741.  
PR 21-OCT-1999; 99US-0160767.  
PR 21-OCT-1999; 99US-0160768.  
PR 21-OCT-1999; 99US-0160770.  
PR 21-OCT-1999; 99US-0160814.  
PR 21-OCT-1999; 99US-0160815.  
PR 22-OCT-1999; 99US-0160980.  
PR 22-OCT-1999; 99US-0160981.  
PR 22-OCT-1999; 99US-0160989.  
PR 25-OCT-1999; 99US-0161404.  
PR 25-OCT-1999; 99US-0161405.  
PR 25-OCT-1999; 99US-0161406.  
PR 26-OCT-1999; 99US-0161359.  
PR 26-OCT-1999; 99US-0161360.  
PR 26-OCT-1999; 99US-0161361.  
PR 28-OCT-1999; 99US-0161920.  
PR 28-OCT-1999; 99US-0161992.  
PR 28-OCT-1999; 99US-0161993.  
PR 29-OCT-1999; 99US-0162142.

Query Match 43.2%; Score 684; DB 21; Length 274;  
Best Local Similarity 49.5%; Pred. No. 6.5e-57;  
Matches 146; Conservative 30; Mismatches 61; Indels 58; Gaps 9;  
QY 1 MGRSPCCERAHNKAGMTKEDQRLINIRVHGCGWRSLPKACGLRCGKSCRLLRWY 60  
||||||| ||||||||| :|||: |||||||||:||||| |||||||||

```

Db      1 mgsrpsccckhntkngawtkeedkllsytkahgegcwrsiprsaglgrcgscrlrtwiny 60.
QY      61 LRPDLKRGNTFEDEDELLIKHSLGNKWSLJAGRLPGFTDNEIKNMYWTHIRKRLISRG 120
Db      61 lrpdlkrgntfeedeallikhsllgnkwsllatrlpgftdneiknmywthvkrkllrkxg 120
QY      121 IDPQTHRPLNQTANTNTATPELDFRNSPTSVSKSSSIKNSLDNPNYNEFOFKSNTDSL 180
Db      121 ldpqthrpnlnektsg-----dsdsdkteedplv-----kllsfqql 158
QY      181 EEPNCTASSGWTDEEOQLHKKOQYGPSNGO--DINLELSIGIVSADSSRVSNANSAE 238
Db      159 ek---lanfg-----derlqkrveyaveercldlnlelrslppwqdklh----- 200
QY      239 SKPKVDNNNFQLEQAMVAKAVCLQWLGFGTSEICRNCON-----SNSNGFYS 287
Db      201 -----dernlrlfgr---vkycrsacrfgfgngkec-scnnvkcgteedassssys 245

RESULT 12
AAE02485
ID      AAE02485 standard; Protein: 274 AA.
XX
AC      AAE02485;
XX
DT      10-AUG-2001 (first entry)
XX
DE      Arabidopsis thaliana transcription factor G664 homologue, G197.
XX
KW      Transcription factor; environmental stress tolerance; gene therapy;
KW      plant structure; plant development.
XX
OS      Arabidopsis thaliana.
XX
PH      Key Location/Qualifiers
FT      Domain 14..119
FT      /note="Conserved domain"
XX
PN      WO200136598-A1.
XX
PD      25-MAY-2001.
XX
PF      14-NOV-2000; 2000MO-US31458.
XX
PR      17-NOV-1999; 99US-0166238.
PR      17-APR-2000; 2000US-0197899.
PR      22-AUG-2000; 2000US-0227439.
XX
XX      (MEND-) MENDEL BIOTECHNOLOGY INC.
PA      (PINE/) PINEDA O.
PA      (YUGC/) YU G.
PA      (CREE/) CREELMAN R.
PA      (RIEC/) RIECHMANN J L.
PA      (HEAR/) HEARD J.
PA      (RATC/) RATCLIFFE O.
PA      (REUB/) REUBER L.
PA      (KEDD/) KEDDIE J.
XX
PI      Pineda O, Yu G, Creelman R, Riechmann JL, Heard J, Ratcliffe O,
PI      Reuber L, Keddie J;
XX
DR      WPI: 2001-336000/35.
XX      N-PSDB: AAD06465.
XX
PT      Nucleic acids encoding plant transcription factor polypeptides, useful
PT      for altering the environmental stress tolerance characteristics of
PT      plants
XX
PS      Claim 4; Page 101-102; 116pp; English.
XX
CC      The present amino acid sequence is Arabidopsis thaliana
CC      transcription factor homologue. This novel transcription factor
CC      is useful for modifying a plant's phenotype in desirable ways, such as

```

```

CC      modifying a plants environmental stress. The transcription factor is
CC      encoded by environmental stress tolerance gene derived from
CC      Arabidopsis thaliana. The transcription factors and the genes encoding
CC      them are used to alter the structure and developmental characteristics
CC      of plants such as soybean, wheat, corn, potato, cotton, rice, oilseed
CC      rape, sunflower, alfalfa, sugarcane, turf, banana, blackberry, papaya,
CC      strawberry, raspberry, cantaloupe, carrot, cauliflower, coffee, onion,
CC      cucumber, eggplant, grapes, honey dew, lettuce, mango, melon, blueberry,
CC      pears, peppers, pineapple, spinach, squash, sweet corn, tobacco, tomato,
CC      watermelon, roseaceous fruits and/or vegetable brassicas. These sequences
CC      are also used for modifying traits associated with environmental stress
CC      tolerance, such as freezing, chilling, heat, drought, water saturation,
CC      salt, photoconditions, radiation and ozone. The transcription factors
CC      are used in gene therapy.
XX
SQ      Sequence 274 AA:

Query Match 43.2%; Score 684; DB 22; Length 274;
Best Local Similarity 49.5%; Pred. No. 6, 5e-57;
Matches 146; Conservative 30; Mismatches 61; Indels 58; Gaps 9;

QY      1 MGRSPCCCKAHNTKNGAWTKEEDORLNIYRVHGGCWRSLPKAGLRCGSCRLRTWINY 60
Db      1 mgsrpsccckhntkngawtkeedkllsytkahgegcwrsiprsaglgrcgscrlrtwiny 60
QY      61 LRPDLKRGNTFEDEDELLIKHSLGNKWSLJAGRLPGFTDNEIKNMYWTHIRKRLISRG 120
Db      61 lrpdlkrgntfeedeallikhsllgnkwsllatrlpgftdneiknmywthvkrkllrkxg 120
QY      121 IDPQTHRPLNQTANTNTATPELDFRNSPTSVSKSSSIKNSLDNPNYNEFOFKSNTDSL 180
Db      121 ldpqthrpnlnektsg-----dsdsdkteedplv-----kllsfqql 158
QY      181 EEPNCTASSGWTDEEOQLHKKOQYGPSNGO--DINLELSIGIVSADSSRVSNANSAE 238
Db      159 ek---lanfg-----derlqkrveyaveercldlnlelrslppwqdklh----- 200
QY      239 SKPKVDNNNFQLEQAMVAKAVCLQWLGFGTSEICRNCON-----SNSNGFYS 287
Db      201 -----dernlrlfgr---vkycrsacrfgfgngkec-scnnvkcgteedassssys 245

RESULT 13
AAG29948
ID      AAG29948 standard; Protein: 269 AA.
XX
AC      AAG29948;
XX
DT      17-OCT-2000 (first entry)
XX
DE      Arabidopsis thaliana protein fragment SEQ ID NO: 35717.
XX
KW      Protein identification; signal transduction pathway; metabolic pathway;
KW      hybridisation assay; genetic mapping; gene expression control; promoter;
KW      termination sequence.
XX
OS      Arabidopsis thaliana.
XX
PI      EP1033405-A2.
XX
PD      06-SEP-2000.
XX
PF      25-FEB-2000; 2000EP-0301439.
XX
PR      25-FEB-1999; 99US-0121825.
PR      05-MAR-1999; 99US-0123180.
PR      09-MAR-1999; 99US-0123548.
PR      23-MAR-1999; 99US-0125788.
PR      25-MAR-1999; 99US-0126264.
PR      29-MAR-1999; 99US-0126785.
PR      01-APR-1999; 99US-0127462.
PR      06-APR-1999; 99US-0128234.

```

PR 08-APR-1999; 99US-0128714.  
PR 16-APR-1999; 99US-0129845.  
PR 19-APR-1999; 99US-0130077.  
PR 21-APR-1999; 99US-0130449.  
PR 23-APR-1999; 99US-0130510.  
PR 23-APR-1999; 99US-0130891.  
PR 28-APR-1999; 99US-0131448.  
PR 30-APR-1999; 99US-0132407.  
PR 04-MAY-1999; 99US-0132484.  
PR 05-MAY-1999; 99US-0132485.  
PR 06-MAY-1999; 99US-0132486.  
PR 07-MAY-1999; 99US-0132863.  
PR 11-MAY-1999; 99US-0134356.  
PR 14-MAY-1999; 99US-0134318.  
PR 14-MAY-1999; 99US-0134219.  
PR 14-MAY-1999; 99US-0134221.  
PR 18-MAY-1999; 99US-0134768.  
PR 19-MAY-1999; 99US-0134841.  
PR 20-MAY-1999; 99US-0135124.  
PR 21-MAY-1999; 99US-0135353.  
PR 24-MAY-1999; 99US-0135629.  
PR 25-MAY-1999; 99US-0136021.  
PR 27-MAY-1999; 99US-0136392.  
PR 28-MAY-1999; 99US-0136782.  
PR 01-JUN-1999; 99US-0137222.  
PR 03-JUN-1999; 99US-0137528.  
PR 04-JUN-1999; 99US-0137502.  
PR 07-JUN-1999; 99US-0137724.  
PR 08-JUN-1999; 99US-0138094.  
PR 10-JUN-1999; 99US-0138540.  
PR 10-JUN-1999; 99US-0138847.  
PR 14-JUN-1999; 99US-0139119.  
PR 16-JUN-1999; 99US-0139452.  
PR 17-JUN-1999; 99US-0139453.  
PR 18-JUN-1999; 99US-0139454.  
PR 18-JUN-1999; 99US-0139455.  
PR 18-JUN-1999; 99US-0139456.  
PR 18-JUN-1999; 99US-0139457.  
PR 18-JUN-1999; 99US-0139458.  
PR 18-JUN-1999; 99US-0139459.  
PR 18-JUN-1999; 99US-0139460.  
PR 18-JUN-1999; 99US-0139461.  
PR 18-JUN-1999; 99US-0139462.  
PR 18-JUN-1999; 99US-0139463.  
PR 18-JUN-1999; 99US-0139464.  
PR 18-JUN-1999; 99US-0139750.  
PR 18-JUN-1999; 99US-0139763.  
PR 21-JUN-1999; 99US-0139817.  
PR 22-JUN-1999; 99US-0139819.  
PR 23-JUN-1999; 99US-0140353.  
PR 23-JUN-1999; 99US-0140354.  
PR 24-JUN-1999; 99US-0140695.  
PR 28-JUN-1999; 99US-0140823.  
PR 29-JUN-1999; 99US-0140991.  
PR 30-JUN-1999; 99US-0141287.  
PR 01-JUL-1999; 99US-0141842.  
PR 02-JUL-1999; 99US-0142154.  
PR 02-JUL-1999; 99US-0142055.  
PR 06-JUL-1999; 99US-0142380.  
PR 08-JUL-1999; 99US-0142803.  
PR 09-JUL-1999; 99US-0142920.  
PR 12-JUL-1999; 99US-0142977.  
PR 13-JUL-1999; 99US-0143542.  
PR 14-JUL-1999; 99US-0143624.  
PR 15-JUL-1999; 99US-0144005.  
PR 16-JUL-1999; 99US-0144085.  
PR 16-JUL-1999; 99US-0144086.  
PR 19-JUL-1999; 99US-0144325.  
PR 19-JUL-1999; 99US-0144331.  
PR 19-JUL-1999; 99US-0144332.

PR 19-JUL-1999; 99US-0144333.  
PR 19-JUL-1999; 99US-0144334.  
PR 19-JUL-1999; 99US-0144335.  
PR 20-JUL-1999; 99US-0144352.  
PR 20-JUL-1999; 99US-0144632.  
PR 20-JUL-1999; 99US-0144864.  
PR 21-JUL-1999; 99US-0144814.  
PR 21-JUL-1999; 99US-0145086.  
PR 21-JUL-1999; 99US-0145088.  
PR 22-JUL-1999; 99US-0145089.  
PR 22-JUL-1999; 99US-0145089.  
PR 22-JUL-1999; 99US-0145192.  
PR 23-JUL-1999; 99US-0145218.  
PR 23-JUL-1999; 99US-0145218.  
PR 23-JUL-1999; 99US-0145224.  
PR 26-JUL-1999; 99US-0145276.  
PR 27-JUL-1999; 99US-0145912.  
PR 27-JUL-1999; 99US-0145918.  
PR 27-JUL-1999; 99US-0145919.  
PR 28-JUL-1999; 99US-0145951.  
PR 02-AUG-1999; 99US-0146386.  
PR 02-AUG-1999; 99US-0146388.  
PR 02-AUG-1999; 99US-0146388.  
PR 03-AUG-1999; 99US-0147038.  
PR 04-AUG-1999; 99US-0147204.  
PR 05-AUG-1999; 99US-0147302.  
PR 05-AUG-1999; 99US-0147192.  
PR 05-AUG-1999; 99US-0147260.  
PR 06-AUG-1999; 99US-0147303.  
PR 06-AUG-1999; 99US-0147416.  
PR 09-AUG-1999; 99US-0147493.  
PR 09-AUG-1999; 99US-0147935.  
PR 10-AUG-1999; 99US-0148171.  
PR 11-AUG-1999; 99US-0148319.  
PR 12-AUG-1999; 99US-0148341.  
PR 13-AUG-1999; 99US-0148565.  
PR 13-AUG-1999; 99US-0148684.  
PR 16-AUG-1999; 99US-0149368.  
PR 17-AUG-1999; 99US-0149175.  
PR 18-AUG-1999; 99US-0149426.  
PR 20-AUG-1999; 99US-0149722.  
PR 20-AUG-1999; 99US-0149723.  
PR 20-AUG-1999; 99US-0149929.  
PR 23-AUG-1999; 99US-0149902.  
PR 23-AUG-1999; 99US-0149930.  
PR 25-AUG-1999; 99US-0150566.  
PR 26-AUG-1999; 99US-0150884.  
PR 27-AUG-1999; 99US-0151065.  
PR 27-AUG-1999; 99US-0151066.  
PR 27-AUG-1999; 99US-0151080.  
PR 30-AUG-1999; 99US-0151303.  
PR 31-AUG-1999; 99US-0151330.  
PR 01-SEP-1999; 99US-0151930.  
PR 07-SEP-1999; 99US-0152363.  
PR 10-SEP-1999; 99US-0153070.  
PR 13-SEP-1999; 99US-0153758.  
PR 13-SEP-1999; 99US-0154018.  
PR 16-SEP-1999; 99US-0154039.  
PR 20-SEP-1999; 99US-0154779.  
PR 22-SEP-1999; 99US-0155139.  
PR 23-SEP-1999; 99US-0155486.  
PR 24-SEP-1999; 99US-0155659.  
PR 28-SEP-1999; 99US-0156458.  
PR 29-SEP-1999; 99US-0156596.  
PR 04-OCT-1999; 99US-0157117.  
PR 05-OCT-1999; 99US-0157753.  
PR 06-OCT-1999; 99US-0157865.  
PR 07-OCT-1999; 99US-0158029.  
PR 08-OCT-1999; 99US-0158232.  
PR 12-OCT-1999; 99US-0158369.  
PR 13-OCT-1999; 99US-0159293.  
PR 13-OCT-1999; 99US-0159294.

```

PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

```

```

Query Match 43.0%; Score 681.5; DB 21; Length 269;
Best Local Similarity 48.8%; Pred. No. 1,1e-56;
Matches 142; Conservative 35; Mismatches 59; Indels 55; Gaps 8;

```

```

OY 1 MGRSPCCERKHTNGKAMTEKEDORLINTIRVHGECWMSLPKAGLLRCGSCRRLRWNTY 60
DB 1 mgrspccerhnmkgawtkeederlvyslshgscwrsldpraagllrcgkscriirwiny 60
OY 61 LRPDLKRGNTFEEDDELIIKHSLLGNKWSLIAGRLPGRTDNEIKYWNTHIRKRLISRG 120
DB 61 lrpdlkrgntfthdedelilkhslgnkwsliagrlpgrtneikynwthirkrliskg 120
OY 121 IDPQTHRLPNOQTANTNTVTAPELDFRNSPTSVSKSSIKNPSLDFNYNEQFKSNTDSL 180
DB 121 idpqrthrlpnoqtantntvtapelefrnsptsvskssiknpsldfnyneofksntdsl 180
OY 181 EEPNCRASSGNTDEO---QEOHLKKQOYGPBSNODINLELSIGIVSADSSRVSNANSA 237
DB 163 gkqnykylnglvckeervvveekl-----gpdlnelelrisppwngqreistctas 212
OY 238 ESKPKVDNNNFOLFEOAMVAKAVCLCWLGFGESEICRNCNSNSNGFYSY 288
DB 213 rf-----ymendmec-----ssetvk-cqlenssrl-sy 239

```

```

RESULT 14
AAE02486
ID AAE02486 standard; Protein; 269 AA.
XX
AC AAE02486;
XX
DT 10-AUG-2001 (first entry)
XX
DE Arabidopsis thaliana transcription factor G664 homologue. G255.
XX
KW Transcription factor; environmental stress tolerance; gene therapy;
XX
KW plant structure; plant development.
XX
OS Arabidopsis thaliana.
XX
FH Key Location/Qualifiers
FT Domain 14..115
FT /note="Conserved domain"
XX
PN WO200136598-A1.
XX

```

```

PD 25-MAY-2001.
XX
PF 14-NOV-2000; 2000WO-US31458.
XX
PR 17-NOV-1999; 99US-0166228.
PR 17-APR-2000; 2000US-0197899.
PR 22-AUG-2000; 2000US-0227439.
XX
PA (MEND-) MENDEL BIOTECHNOLOGY INC.
PA (PINE/) PINEDA O.
PA (YUG/) YU G.
PA (CREE/) CREELMAN R.
PA (RIEC/) RIECHMANN J L.
PA (HEAR/) HEARD J.
PA (RATC/) RATCLIFFE O.
PA (REUB/) REUBER L.
PA (KEDD/) KEDDIE J.
XX
PI Pineda O, Yu G, Creelman R, Riechmann JL, Heard J, Ratcliffe O,
PI Reuber L, Keddie J;
XX
DR WPI; 2001-336000/35.
DR N-PSDB; AAD06466.
XX
PT Nucleic acids encoding plant transcription factor polypeptides, useful
PT for altering the environmental stress tolerance characteristics of
PT plants.
XX
XX
XX Claim 4; Page 103-104; 116pp; English.
XX

```

The present amino acid sequence is Arabidopsis thaliana transcription factor homologue. This novel transcription factor is useful for modifying a plant's phenotype in desirable ways, such as modifying a plant's environmental stress. The transcription factor is encoded by environmental stress tolerance gene derived from Arabidopsis thaliana. The transcription factors and the genes encoding them are used to alter the structure and developmental characteristics of plants such as soybean, wheat, corn, potato, cotton, rice, oilseed rape, sunflower, alfalfa, sugarcane, turf, banana, blackberry, papaya, strawberry, raspberry, cantaloupe, carrot, cauliflower, coffee, onion, cucumber, eggplant, grapes, honey dew, lettuce, mango, melon, blueberry, peas, peppers, pineapple, spinach, squash, sweet corn, tobacco, tomato, watermelon, rosehip, fruits and/or vegetable brassicas. These sequences are also used for modifying traits associated with environmental stress tolerance, such as freezing, chilling, heat, drought, water saturation, salt, photoconditions, radiation and ozone. The transcription factors are used in gene therapy.

```

Query Match 43.0%; Score 681.5; DB 22; Length 269;
Best Local Similarity 48.8%; Pred. No. 1,1e-56;
Matches 142; Conservative 35; Mismatches 59; Indels 55; Gaps 8;

```

```

OY 1 MGRSPCCERKHTNGKAMTEKEDORLINTIRVHGECWMSLPKAGLLRCGSCRRLRWNTY 60
DB 1 mgrspccerhnmkgawtkeederlvyslshgscwrsldpraagllrcgkscriirwiny 60
OY 61 LRPDLKRGNTFEEDDELIIKHSLLGNKWSLIAGRLPGRTDNEIKYWNTHIRKRLISRG 120
DB 61 lrpdlkrgntfthdedelilkhslgnkwsliagrlpgrtneikynwthirkrliskg 120
OY 121 IDPQTHRLPNOQTANTNTVTAPELDFRNSPTSVSKSSIKNPSLDFNYNEQFKSNTDSL 180
DB 121 idpqrthrlpnoqtantntvtapelefrnsptsvskssiknpsldfnyneofksntdsl 180
OY 181 EEPNCRASSGNTDEO---QEOHLKKQOYGPBSNODINLELSIGIVSADSSRVSNANSA 237
DB 163 gkqnykylnglvckeervvveekl-----gpdlnelelrisppwngqreistctas 212
OY 238 ESKPKVDNNNFOLFEOAMVAKAVCLCWLGFGESEICRNCNSNSNGFYSY 288

```

Db 213 rf-----ymendmec-----ssetvk-cqltenssl-sy 239

RESULT 15  
AAG29947  
ID AAG29947 standard; Protein; 278 AA.  
XX AAG29947;  
AC  
XX  
DT 17-OCT-2000 (first entry)  
XX  
DE Arabidopsis thaliana protein fragment SEQ ID NO: 35716.  
XX  
KW Protein identification; signal transduction pathway; metabolic pathway;  
hybridisation assay; genetic mapping; gene expression control; promoter;  
termination sequence.  
XX  
OS Arabidopsis thaliana.  
XX  
PN EP1033405-A2.  
XX  
PD 06-SEP-2000.  
XX  
PE 25-FEB-2000; 2000EP-0301439.  
XX  
PR 25-FEB-1999; 99US-0121825.  
PR 05-MAR-1999; 99US-0123180.  
PR 09-MAR-1999; 99US-0123548.  
PR 23-MAR-1999; 99US-0125788.  
PR 25-MAR-1999; 99US-0126264.  
PR 29-MAR-1999; 99US-0126785.  
PR 01-APR-1999; 99US-0128234.  
PR 06-APR-1999; 99US-0128714.  
PR 16-APR-1999; 99US-0129845.  
PR 19-APR-1999; 99US-0130077.  
PR 21-APR-1999; 99US-0130449.  
PR 23-APR-1999; 99US-0130510.  
PR 28-APR-1999; 99US-0130691.  
PR 30-APR-1999; 99US-0131448.  
PR 30-APR-1999; 99US-0132048.  
PR 04-MAY-1999; 99US-0132407.  
PR 05-MAY-1999; 99US-0132485.  
PR 06-MAY-1999; 99US-0132486.  
PR 07-MAY-1999; 99US-0132487.  
PR 11-MAY-1999; 99US-0132863.  
PR 14-MAY-1999; 99US-0134256.  
PR 14-MAY-1999; 99US-0134219.  
PR 14-MAY-1999; 99US-0134221.  
PR 14-MAY-1999; 99US-0134370.  
PR 18-MAY-1999; 99US-0134768.  
PR 19-MAY-1999; 99US-0134941.  
PR 20-MAY-1999; 99US-0135124.  
PR 21-MAY-1999; 99US-0135353.  
PR 25-MAY-1999; 99US-0135629.  
PR 25-MAY-1999; 99US-0136021.  
PR 27-MAY-1999; 99US-0136392.  
PR 28-MAY-1999; 99US-0136782.  
PR 01-JUN-1999; 99US-0137222.  
PR 03-JUN-1999; 99US-0137528.  
PR 04-JUN-1999; 99US-0137562.  
PR 07-JUN-1999; 99US-0137724.  
PR 08-JUN-1999; 99US-0138094.  
PR 10-JUN-1999; 99US-0138540.  
PR 10-JUN-1999; 99US-0138847.  
PR 14-JUN-1999; 99US-0139119.  
PR 16-JUN-1999; 99US-0139452.  
PR 16-JUN-1999; 99US-0139453.  
PR 17-JUN-1999; 99US-0139452.  
PR 18-JUN-1999; 99US-0139454.  
PR 18-JUN-1999; 99US-0139455.

PR 18-JUN-1999; 99US-0139456.  
PR 18-JUN-1999; 99US-0139457.  
PR 18-JUN-1999; 99US-0139458.  
PR 18-JUN-1999; 99US-0139459.  
PR 18-JUN-1999; 99US-0139460.  
PR 18-JUN-1999; 99US-0139461.  
PR 18-JUN-1999; 99US-0139462.  
PR 18-JUN-1999; 99US-0139463.  
PR 18-JUN-1999; 99US-0139750.  
PR 18-JUN-1999; 99US-0139763.  
PR 21-JUN-1999; 99US-0139817.  
PR 22-JUN-1999; 99US-0139899.  
PR 23-JUN-1999; 99US-0140353.  
PR 23-JUN-1999; 99US-0140354.  
PR 24-JUN-1999; 99US-0140695.  
PR 28-JUN-1999; 99US-0140823.  
PR 29-JUN-1999; 99US-0140921.  
PR 30-JUN-1999; 99US-0141287.  
PR 01-JUL-1999; 99US-0141642.  
PR 01-JUL-1999; 99US-0142154.  
PR 02-JUL-1999; 99US-0142055.  
PR 06-JUL-1999; 99US-0142390.  
PR 08-JUL-1999; 99US-0142803.  
PR 09-JUL-1999; 99US-0142920.  
PR 12-JUL-1999; 99US-0142977.  
PR 13-JUL-1999; 99US-0143542.  
PR 14-JUL-1999; 99US-0143624.  
PR 15-JUL-1999; 99US-0144005.  
PR 16-JUL-1999; 99US-0144085.  
PR 16-JUL-1999; 99US-0144086.  
PR 19-JUL-1999; 99US-0144325.  
PR 19-JUL-1999; 99US-0144331.  
PR 19-JUL-1999; 99US-0144332.  
PR 19-JUL-1999; 99US-0144333.  
PR 19-JUL-1999; 99US-0144334.  
PR 19-JUL-1999; 99US-0144335.  
PR 20-JUL-1999; 99US-0144352.  
PR 20-JUL-1999; 99US-0144632.  
PR 20-JUL-1999; 99US-0144684.  
PR 21-JUL-1999; 99US-0144814.  
PR 21-JUL-1999; 99US-0145086.  
PR 21-JUL-1999; 99US-0145088.  
PR 22-JUL-1999; 99US-0145087.  
PR 22-JUL-1999; 99US-0145087.  
PR 22-JUL-1999; 99US-0145089.  
PR 22-JUL-1999; 99US-0145192.  
PR 23-JUL-1999; 99US-0145145.  
PR 23-JUL-1999; 99US-0145218.  
PR 23-JUL-1999; 99US-0145224.  
PR 26-JUL-1999; 99US-0145276.  
PR 27-JUL-1999; 99US-0145913.  
PR 27-JUL-1999; 99US-0145918.  
PR 27-JUL-1999; 99US-0145919.  
PR 28-JUL-1999; 99US-0145951.  
PR 02-AUG-1999; 99US-0146386.  
PR 02-AUG-1999; 99US-0146388.  
PR 02-AUG-1999; 99US-0146389.  
PR 03-AUG-1999; 99US-0147038.  
PR 04-AUG-1999; 99US-0147204.  
PR 04-AUG-1999; 99US-0147302.  
PR 05-AUG-1999; 99US-0147192.  
PR 05-AUG-1999; 99US-0147260.  
PR 06-AUG-1999; 99US-0147303.  
PR 06-AUG-1999; 99US-0147416.  
PR 09-AUG-1999; 99US-0147493.  
PR 09-AUG-1999; 99US-0147935.  
PR 10-AUG-1999; 99US-0148171.  
PR 11-AUG-1999; 99US-0148319.  
PR 12-AUG-1999; 99US-0148341.  
PR 13-AUG-1999; 99US-0148565.  
PR 13-AUG-1999; 99US-0148684.  
PR 16-AUG-1999; 99US-0149368.  
PR 17-AUG-1999; 99US-0149375.



PR 18-AUG-1999; 99US-0149426.  
 PR 20-AUG-1999; 99US-0149722.  
 PR 20-AUG-1999; 99US-0149723.  
 PR 20-AUG-1999; 99US-0149929.  
 PR 23-AUG-1999; 99US-0149902.  
 PR 23-AUG-1999; 99US-0149930.  
 PR 25-AUG-1999; 99US-0150566.  
 PR 26-AUG-1999; 99US-0150884.  
 PR 27-AUG-1999; 99US-0151065.  
 PR 27-AUG-1999; 99US-0151066.  
 PR 27-AUG-1999; 99US-0151080.  
 PR 30-AUG-1999; 99US-0151303.  
 PR 31-AUG-1999; 99US-0151438.  
 PR 01-SEP-1999; 99US-0151930.  
 PR 07-SEP-1999; 99US-0152363.  
 PR 10-SEP-1999; 99US-0153070.  
 PR 13-SEP-1999; 99US-0153758.  
 PR 15-SEP-1999; 99US-0154018.  
 PR 16-SEP-1999; 99US-0154039.  
 PR 20-SEP-1999; 99US-0154779.  
 PR 22-SEP-1999; 99US-0155139.  
 PR 23-SEP-1999; 99US-0155486.  
 PR 24-SEP-1999; 99US-0155659.  
 PR 28-SEP-1999; 99US-0156458.  
 PR 29-SEP-1999; 99US-0156596.  
 PR 04-OCT-1999; 99US-0157117.  
 PR 05-OCT-1999; 99US-0157753.  
 PR 06-OCT-1999; 99US-0157865.  
 PR 07-OCT-1999; 99US-0158029.  
 PR 08-OCT-1999; 99US-0158232.  
 PR 12-OCT-1999; 99US-0158369.  
 PR 13-OCT-1999; 99US-0159293.  
 PR 13-OCT-1999; 99US-0159294.  
 PR 13-OCT-1999; 99US-0159295.  
 PR 14-OCT-1999; 99US-0159329.  
 PR 14-OCT-1999; 99US-0159330.  
 PR 14-OCT-1999; 99US-0159331.  
 PR 14-OCT-1999; 99US-0159637.  
 PR 14-OCT-1999; 99US-0159638.  
 PR 18-OCT-1999; 99US-0159584.  
 PR 21-OCT-1999; 99US-0160741.  
 PR 21-OCT-1999; 99US-0160767.  
 PR 21-OCT-1999; 99US-0160768.  
 PR 21-OCT-1999; 99US-0160770.  
 PR 21-OCT-1999; 99US-0160814.  
 PR 21-OCT-1999; 99US-0160815.  
 PR 22-OCT-1999; 99US-0160980.  
 PR 22-OCT-1999; 99US-0160981.  
 PR 22-OCT-1999; 99US-0160989.  
 PR 25-OCT-1999; 99US-0161404.  
 PR 25-OCT-1999; 99US-0161405.  
 PR 25-OCT-1999; 99US-0161406.  
 PR 26-OCT-1999; 99US-0161359.  
 PR 26-OCT-1999; 99US-0161360.  
 PR 26-OCT-1999; 99US-0161361.  
 PR 28-OCT-1999; 99US-0161920.  
 PR 28-OCT-1999; 99US-0161992.  
 PR 28-OCT-1999; 99US-0161993.  
 PR 29-OCT-1999; 99US-0162142.

Query Match 43.0%; Score 681.5; DB 21; Length 278;  
 Best Local Similarity 48.8%; Pred. No. 1.2e-56;  
 Matches 142; Conservative 35; Mismatches 59; Indels 55; Gaps 8;

QY 1 MGRSPCEKANTNGATWKEEDQRLINIRVHGEBCWRLPKAAGLLRGGKSCRLRWINY 60  
 DB 10 mgyrpscekekmnkgawckederlveyiknqgcwrsipraagllrcqkscrlrwny 69  
 QY 61 LRPLKGNFTFEEDELIKLSLGNKWSLITAGRLPGRTDNEIKNYWTHIKRKLISRG 120  
 DB 70 lrplkrgnftfededelikkhsllgnkwsllaarlpgrtdneiknywthikrkliskg 129  
 QY 121 IDPOTHRPLNOTANTNTVATPTELDFRNSPTSVSKSSSIKNPSLDENYNEFOFKSNTDSL 180

DB 130 ldpnthrginea-----kisdllkttdqivkdvsvfkfee----tdksqd 171  
 QY 181 EEPNCTASSGATTDEO---QEOHLKROOYGPSNGODINLEISIGIVSADSSRVSNANSA 237  
 DB 172 qkqmkylrnglsvckeervvveeki-----gpdlnlelrtrispwqqrrelstctas 221  
 QY 238 ESKPKVDNNNEQFLQAMVAKAVCLCMQDGFSTSEICRCNCONSNNGFYX 208  
 DB 222 rfy-----ymendnec-----ssetvkc-qctensssl-sy 248

Search completed: January 23, 2002, 21:11:32  
 Job time: 3050 sec

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 23, 2002, 20:22:32 ; Search time 64.93 seconds  
(without alignments)  
101.894 Million cell updates/sec

Title: US-09-453-387A-2  
Perfect score: 1584  
Sequence: 1 MGRSPCCCKAKATNKAWTKE.....RNCONSNNGFYSCRPDLS 294

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 212252 seqs, 22503292 residues

Total number of hits satisfying chosen parameters: 212252

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database : Issued\_Patents\_AA:\*  
1: /cgn2\_6/ptodata/2/1aa/5A.COMB.pep:\*  
2: /cgn2\_6/ptodata/2/1aa/5B.COMB.pep:\*  
3: /cgn2\_6/ptodata/2/1aa/5A.COMB.pep:\*  
4: /cgn2\_6/ptodata/2/1aa/5B.COMB.pep:\*  
5: /cgn2\_6/ptodata/2/1aa/PTUS.COMB.pep:\*  
6: /cgn2\_6/ptodata/2/1aa/backfiles1.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	493	31.1	278	2	US-08-722-626B-2
2	394.5	24.9	553	4	US-08-997-251-4
3	393.5	24.8	553	4	US-08-997-251-2
4	282.5	17.8	156	4	US-08-928-941D-4
5	282.5	17.8	156	4	US-08-928-941D-36
6	282.5	17.8	156	4	US-09-280-590A-4
7	282.5	17.8	156	4	US-09-280-590A-46
8	167	10.5	51	1	US-08-519-103-17
9	167	10.5	51	1	US-09-018-635-17
10	163	10.3	51	1	US-08-519-103-16
11	163	10.3	51	1	US-08-519-103-16
12	163	10.3	51	1	US-09-018-635-16
13	163	10.3	51	4	US-09-018-635-19
14	160	10.1	51	1	US-08-519-103-18
15	160	10.1	51	4	US-09-018-635-18
16	137	8.6	802	4	US-08-928-941D-29
17	136.5	8.6	760	4	US-09-280-590A-29
18	136.5	8.6	760	4	US-09-280-590A-29
19	132.5	8.4	94	2	US-08-814-030-1
20	127	8.0	761	4	US-08-928-941D-1
21	127	8.0	761	4	US-09-280-590A-1
22	123.5	7.7	372	4	US-08-928-941D-16
23	123.5	7.7	372	4	US-09-280-590A-16
24	121.5	7.7	52	1	US-08-519-103-13
25	121.5	7.7	52	1	US-08-519-103-14
26	121.5	7.7	52	4	US-09-018-635-13
27	121.5	7.7	52	4	US-09-018-635-14

28	117.5	7.4	50	4	US-09-156-316-4	Sequence 4, Appl
29	116.5	7.4	50	4	US-09-156-316-5	Sequence 5, Appl
30	116.5	7.4	50	4	US-09-156-316-6	Sequence 6, Appl
31	112.5	7.1	169	4	US-08-928-941D-35	Sequence 35, Appl
32	112.5	7.1	169	4	US-09-280-590A-45	Sequence 45, Appl
33	101	6.4	711	3	US-08-613-009A-21	Sequence 21, Appl
34	92	5.8	412	2	US-08-741-134-2	Sequence 22, Appl
35	90.5	5.7	1202	1	US-08-425-061-22	Sequence 22, Appl
36	90.5	5.7	1202	2	US-08-825-886-22	Sequence 22, Appl
37	90.5	5.7	1363	1	US-08-425-061-23	Sequence 23, Appl
38	90.5	5.7	1363	2	US-08-825-886-23	Sequence 23, Appl
39	90.5	5.7	1852	1	US-08-425-061-24	Sequence 24, Appl
40	90.5	5.7	1852	2	US-08-825-886-24	Sequence 24, Appl
41	90.5	5.7	1863	1	US-08-425-061-16	Sequence 16, Appl
42	90.5	5.7	1863	1	US-08-593-591-2	Sequence 2, Appl
43	90.5	5.7	1863	1	US-08-480-784-2	Sequence 2, Appl
44	90.5	5.7	1863	1	US-08-483-553-2	Sequence 2, Appl
45	90.5	5.7	1863	2	US-08-603-753D-2	Sequence 2, Appl

## ALIGNMENTS

RESULT 1  
US-08-722-626B-2  
Sequence 2, Application US/08722626B  
Patent No. 5939601  
GENERAL INFORMATION:  
APPLICANT: Yang, Yinnong  
TITLE OF INVENTION: NOVEL GENES ASSOCIATED WITH ENHANCED  
DISEASE RESISTANCE IN PLANTS  
NUMBER OF SEQUENCES: 9  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Dann, Dorfman, Herrell and Skillman  
STREET: 1601 Market Street Suite 720  
CITY: Philadelphia  
STATE: PA  
COUNTRY: USA  
ZIP: 19103-2307  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FASTED Version 1.5  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/722,626B  
FILING DATE: 27-SEP-1996  
CLASSIFICATION: 800  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER:  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Pat Hagan  
REGISTRATION NUMBER: 27,643  
REFERENCE/DOCKET NUMBER: 97-0010  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 215 563-4100  
TELEFAX: 215 563-4044  
TELEX:  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 278 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
FRAGMENT TYPE: internal  
ORIGINAL SOURCE:  
US-08-722-626B-2

Query Match 31.1%; Score 493; DB 2; Length 278;  
Best Local Similarity 38.8%; Pred. No. 5.5e-44;  
Matches 111; Conservative 40; Mismatches 75; Indels 60; Gaps 7;

QY 1 MGRSPCKEAKHTNKGAWTKEEDQRLINIRVHGEGRSLPKAAGLRGKSCRLRWNTY 60  
DB 1 MVRAPCCCKMGKAGKPMWEEDQILISFIQTNGHNMWRALPKAAGLRGKSCRLRWNTY 60  
QY 61 LRPDLKRGFTFEEDDLIKLHSLGNKSLINGRLPGRDNETKNWNTNHRKRLISRG 120  
DB 61 LRPDLKRGFTFEEDDLIKLHSLGNKSLINGRLPGRDNETKNWNTNHRKRLISRG 118  
QY 121 IDPQHRPLNGANTNTVAPTELDPRNSP--TSVSKSSIKNPSLDENYNEFOKSNMTD 178  
DB 119 YKPPONSKRHSKSKNDKSGFTTSSNSDITITNKHIDSVLPN----- 167  
QY 179 SLEPNCRTASSGAT---DEEQEQLHKQOYGPSPNGODINLELSIGIVSADSSRVSA 234  
DB 168 ---SPOISSSTMTSTYTLVDHDMYI--KQEV-----M 196  
QY 235 NSAESKPKVD-----NNNEQLEQAMAKAVCLCWLQGFCT 270  
DB 197 ESSEYFPEIDESEFTMDLFTDNNMSSSTDHVMANDELQVQLPSS 242

RESULT 2  
US-08-997-251-4  
; Sequence 4, Application US/08997251  
; Patent No. 6271440

GENERAL INFORMATION:  
APPLICANT: GUBLER, FRANZ J.  
APPLICANT: JACOBSEN, JOHN V.  
TITLE OF INVENTION: PLANT REGULATORY PROTEINS III  
NUMBER OF SEQUENCES: 17  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Greenlee, Winner and Sullivan, P.C.  
STREET: 5370 Manhattan Boulevard  
CITY: Boulder  
STATE: CO  
COUNTRY: US  
ZIP: 80303  
COMPUTER READABLE FORM:  
MEDIUM TYPE: floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/997,251  
FILING DATE: 23-DEC-1997  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: WO A/96/00383  
FILING DATE: 21-JUN-1996  
APPLICATION DATA:  
APPLICATION NUMBER: AU P/64470/95  
FILING DATE: 09-NOV-1995  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: AU P/3779/95  
FILING DATE: 23-JUN-1995  
ATTORNEY/AGENT INFORMATION:  
NAME: Winner, Ellen P.  
REGISTRATION NUMBER: 28,547  
REFERENCE/DOCKET NUMBER: 110-97  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (303) 499-8080  
TELEFAX: (303) 499-8089  
INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 553 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein

US-08-997-251-4

Query Match 24.9%; Score 394.5; DB 4; Length 553;  
Best Local Similarity 35.4%; Pred. No. 4.2e-33;  
Matches 107; Conservative 49; Mismatches 91; Indels 55; Gaps 13;

QY 14 KGAMTEEDQRLINIRVHGEGRSLPKAAGLRGKSCRLRWNTYLRPDLKRGFTFE 73  
DB 42 KGPWTSADPAIDLVDYVKKHGEGRNNAVQKNTGLFCGKSCRLRWNTNHRKRLISRG 101  
QY 74 EDLILIKLHSLGNKSLINGRLPGRDNETKNWNTNHRKRLISRGIDPQHRPLN 133  
DB 102 EERLLIQLHSMKMGKAMAHLPGRDNETKNWNTNHRKRLISRGIDPQHRPLN 160  
QY 134 NRTVTVAPTELDP---RNSPTSVSKSSIKNPSLDENYNEFOKSNMTDLE-EPNCTASS 189  
DB 161 NEDQOCS---DFDGGENLSNDLNLNGLYLP--DFTCD--NFANSEALPYAPHLISAVS 213  
QY 190 ---GWTDEEQLHKQOYGPSPNGODINLELSIGIVSADSSRVSAESKPKVD 244  
DB 214 ISNLLGSEFASKSCSEFMDVQNTG-----NLKQSDGLVLPGLSDPTINGVIS- 261  
QY 245 ---NNNEQLEQAMAKAVCLCWLQGF-----GTSET-----CRNCONSNGFTSYC 289  
DB 262 QFSNDESKLQA-----VGFYDLHFNANSTSKITAPGALNGSHAFNGFN 310  
QY 290 RP 291  
DB 311 RP 312

RESULT 3  
US-08-997-251-2  
; Sequence 2, Application US/08997251  
; Patent No. 6271440

GENERAL INFORMATION:  
APPLICANT: GUBLER, FRANZ J.  
APPLICANT: JACOBSEN, JOHN V.  
TITLE OF INVENTION: PLANT REGULATORY PROTEINS III  
NUMBER OF SEQUENCES: 17  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Greenlee, Winner and Sullivan, P.C.  
STREET: 5370 Manhattan Boulevard  
CITY: Boulder  
STATE: CO  
COUNTRY: US  
ZIP: 80303  
COMPUTER READABLE FORM:  
MEDIUM TYPE: floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/997,251  
FILING DATE: 23-DEC-1997  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: WO A/96/00383  
FILING DATE: 21-JUN-1996  
APPLICATION DATA:  
APPLICATION NUMBER: AU P/64470/95  
FILING DATE: 09-NOV-1995  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: AU P/3779/95  
FILING DATE: 23-JUN-1995  
ATTORNEY/AGENT INFORMATION:  
NAME: Winner, Ellen P.  
REGISTRATION NUMBER: 28,547  
REFERENCE/DOCKET NUMBER: 110-97  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (303) 499-8080  
TELEFAX: (303) 499-8089

```

US-08-928-941D-36
; Sequence 36, Application US/08928941D
; Patent No. 6180763
; GENERAL INFORMATION:
; APPLICANT: Hirai, Hiroshi
; APPLICANT: Sherf, Charles
; TITLE OF INVENTION: CYCLIN-D BINDING FACTOR, AND USES
; TITLE OF INVENTION: THEREOF
; NUMBER OF SEQUENCES: 36
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: David A. Jackson, Esq.
; STREET: 411 Hackensack Ave, Continental Plaza, 4th
; STREET: Floor
; CITY: Hackensack
; STATE: New Jersey
; COUNTRY: USA
; ZIP: 07601
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/928,941D
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Jackson Esq., David A.
; REGISTRATION NUMBER: 26,742
; REFERENCE/DOCKET NUMBER: 1340-1-002 N CIP
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 201-487-5800
; TELEFAX: 201-943-1684
; INFORMATION FOR SEQ ID NO: 36:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 156 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: YES
; FRAGMENT TYPE:
; ORIGINAL SOURCE:
; ORGANISM: Gallus gallus
; US-08-928-941D-36

Query Match 17.8%; Score 282.5; DB 4; Length 156;
Best Local Similarity 51.0%; Pred No. 4e-22;
Matches 53; Conservative 20; Mismatches 3; Gaps 2;

```

OY 14 KGAWTKEDORLINTYRHWGEGCWSLPR-AAGLLRCGKSCRLWNTYLRPDLKRGNETE 72  
DB 55 KGPWTKEDORVITLVOKIGPKRMSVIAKHKG--RIGKOCREKRNHNLNPEVKTISWTE 112  
OY 73 EDEDLITIKLSLGNKMSLIGRLPGRTDNEIKNYWNTHIKRL 116  
DB 113 EEDRIIYQAHKRILGNRMAEIAKLPLPGRTDIAIKNMNSTMRKV 156

RESULT 6  
US-09-280-590A-4  
Sequence 4, Application US/09280590A  
Patent No. 6303772  
GENERAL INFORMATION:  
APPLICANT: Hirai, Hiroshi  
Sheri, Charles  
Inoue, Kazushi  
Bodner, Sarah M.  
TITLE OF INVENTION: CYCLIN-D BINDING FACTOR, AND USES THEREOF  
NUMBER OF SEQUENCES: 46  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: David A. Jackson, Esq.  
STREET: 411 Hackensack Ave, Continental Plaza, 4th Floor  
CITY: Hackensack  
STATE: New Jersey  
COUNTRY: USA  
ZIP: 07601  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/280,590A  
FILING DATE: 29-Mar-1999  
CLASSIFICATION: <Unknown>  
ATTORNEY/AGENT INFORMATION:  
NAME: Jackson Esq., David A.  
REGISTRATION NUMBER: 26,742  
REFERENCE/DOCKET NUMBER: 1340-1-002 N CP2  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 201-487-5800  
TELEFAX: 201-343-1684  
INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 156 amino acids  
TYPE: amino acid  
STRANDEDNESS: <Unknown>  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
HYPOTHEICAL: NO  
FRAGMENT TYPE: internal  
ORIGINAL SOURCE:  
ORGANISM: Mus musculus  
SEQUENCE DESCRIPTION: SEQ ID NO: 4:  
US-09-280-590A-4

Query Match 17.8%; Score 282.5; DB 4; Length 156;  
Best Local Similarity 51.0%; Pred. No. 4e-22; 28; Indels 3; Gaps 2;  
Matches 53; Conservative 20; Mismatches 28;

OY 14 KGAWTKEDORLINTYRHWGEGCWSLPR-AAGLLRCGKSCRLWNTYLRPDLKRGNETE 72  
DB 55 KGPWTKEDORVITLVOKIGPKRMSVIAKHKG--RIGKOCREKRNHNLNPEVKTISWTE 112  
OY 73 EDEDLITIKLSLGNKMSLIGRLPGRTDNEIKNYWNTHIKRL 116  
DB 113 EEDRIIYQAHKRILGNRMAEIAKLPLPGRTDIAIKNMNSTMRKV 156

RESULT 7  
US-09-280-590A-46  
Sequence 46, Application US/09280590A  
Patent No. 6303772  
GENERAL INFORMATION:  
APPLICANT: Hirai, Hiroshi  
Sheri, Charles  
Inoue, Kazushi  
Bodner, Sarah M.  
TITLE OF INVENTION: CYCLIN-D BINDING FACTOR, AND USES THEREOF  
NUMBER OF SEQUENCES: 46  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: David A. Jackson, Esq.  
STREET: 411 Hackensack Ave, Continental Plaza, 4th Floor  
CITY: Hackensack  
STATE: New Jersey  
COUNTRY: USA  
ZIP: 07601  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/280,590A  
FILING DATE: 29-Mar-1999  
CLASSIFICATION: <Unknown>  
ATTORNEY/AGENT INFORMATION:  
NAME: Jackson Esq., David A.  
REGISTRATION NUMBER: 26,742  
REFERENCE/DOCKET NUMBER: 1340-1-002 N CP2  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 201-487-5800  
TELEFAX: 201-343-1684  
INFORMATION FOR SEQ ID NO: 46:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 156 amino acids  
TYPE: amino acid  
STRANDEDNESS: <Unknown>  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
HYPOTHEICAL: YES  
FRAGMENT TYPE: <Unknown>  
ORIGINAL SOURCE:  
ORGANISM: Gallus gallus  
SEQUENCE DESCRIPTION: SEQ ID NO: 46:  
US-09-280-590A-46

Query Match 17.8%; Score 282.5; DB 4; Length 156;  
Best Local Similarity 51.0%; Pred. No. 4e-22; 28; Indels 3; Gaps 2;  
Matches 53; Conservative 20; Mismatches 28;

OY 14 KGAWTKEDORLINTYRHWGEGCWSLPR-AAGLLRCGKSCRLWNTYLRPDLKRGNETE 72  
DB 55 KGPWTKEDORVITLVOKIGPKRMSVIAKHKG--RIGKOCREKRNHNLNPEVKTISWTE 112  
OY 73 EDEDLITIKLSLGNKMSLIGRLPGRTDNEIKNYWNTHIKRL 116  
DB 113 EEDRIIYQAHKRILGNRMAEIAKLPLPGRTDIAIKNMNSTMRKV 156

RESULT 8  
US-08-519-103-17  
Sequence 17, Application US/08519103  
Patent No. 5733730  
GENERAL INFORMATION:  
APPLICANT: delange, Titia  
TITLE OF INVENTION: TELOMERE REPEAT BINDING FACTOR AND  
TITLE OF INVENTION: DIAGNOSTIC AND THERAPEUTIC USE THEREOF

NUMBER OF SEQUENCES: 21  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: KLAUBER & JACKSON  
 STREET: 411 Hackensack Avenue  
 CITY: Hackensack  
 STATE: New Jersey  
 COUNTRY: USA  
 ZIP: 07601  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: Patentin Release #1.0, Version #1.30  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/519,103  
 FILING DATE: 25-AUG-1995  
 CLASSIFICATION: 435  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Crane-Feury, Sharon E.  
 REGISTRATION NUMBER: 36,113  
 REFERENCE/DOCKET NUMBER: 600-1-142  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: 201-487-5800  
 TELEFAX: 201-343-1684  
 TELEX: 133521  
 INFORMATION FOR SEQ ID NO: 17:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 51 amino acids  
 TYPE: amino acid  
 STRANDEDNESS: single  
 TOPOLOGY: linear  
 MOLECULE TYPE: peptide  
 US-08-519-103-17

Query Match 10.5%; Score 167; DB 1; Length 51;  
 Best Local Similarity 56.9%; Pred. No. 1.1e-10;  
 Matches 29; Conservative 11; Mismatches 11; Indels 0; Gaps 0;

OY 65 LKRGNTEEDDELITIKLSLGNKWSLIAGRLPGRTDNEIKYNWTHIRK 115  
 Db 1 VKKSCWTEEDRITICEAHKVLGNRAEIAKMLPGRTDNVKNHNMSTIKRK 51

RESULT 9  
 US-09-018-635-17  
 Sequence 17, Application US/09018635  
 Patent No. 6297356  
 GENERAL INFORMATION:  
 APPLICANT: de Lange, Titia  
 APPLICANT: Broccoli, Dominique  
 APPLICANT: Smogorzewska, Agata  
 TITLE OF INVENTION: TELOMERE REPEAT BINDING FACTOR AND  
 TITLE OF INVENTION: DIAGNOSTIC AND THERAPEUTIC USE THEREOF  
 NUMBER OF SEQUENCES: 52  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: KLAUBER & JACKSON  
 STREET: 411 Hackensack Avenue  
 CITY: Hackensack  
 STATE: New Jersey  
 COUNTRY: USA  
 ZIP: 07601  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: Patentin Release #1.0, Version #1.30  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/09/018,635  
 FILING DATE:  
 CLASSIFICATION:  
 ATTORNEY/AGENT INFORMATION:  
 NAME: David A. Jackson

REGISTRATION NUMBER: 26,742  
 REFERENCE/DOCKET NUMBER: 600-1-142 CIP1  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: 201-487-5800  
 TELEFAX: 201-343-1684  
 TELEX: 133521  
 INFORMATION FOR SEQ ID NO: 17:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 51 amino acids  
 TYPE: amino acid  
 STRANDEDNESS: single  
 TOPOLOGY: linear  
 MOLECULE TYPE: peptide  
 US-09-018-635-17

Query Match 10.5%; Score 167; DB 4; Length 51;  
 Best Local Similarity 56.9%; Pred. No. 1.1e-10;  
 Matches 29; Conservative 11; Mismatches 11; Indels 0; Gaps 0;

OY 65 LKRGNTEEDDELITIKLSLGNKWSLIAGRLPGRTDNEIKYNWTHIRK 115  
 Db 1 VKKSCWTEEDRITICEAHKVLGNRAEIAKMLPGRTDNVKNHNMSTIKRK 51

RESULT 10  
 US-08-519-103-16  
 Sequence 16, Application US/08519103  
 Patent No. 5733730  
 GENERAL INFORMATION:  
 APPLICANT: deLange, Titia  
 TITLE OF INVENTION: TELOMERE REPEAT BINDING FACTOR AND  
 TITLE OF INVENTION: DIAGNOSTIC AND THERAPEUTIC USE THEREOF  
 NUMBER OF SEQUENCES: 21  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: KLAUBER & JACKSON  
 STREET: 411 Hackensack Avenue  
 CITY: Hackensack  
 STATE: New Jersey  
 COUNTRY: USA  
 ZIP: 07601  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 OPERATING SYSTEM: IBM PC compatible  
 SOFTWARE: Patentin Release #1.0, Version #1.30  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/519,103  
 FILING DATE: 25-AUG-1995  
 CLASSIFICATION: 435  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Crane-Feury, Sharon E.  
 REGISTRATION NUMBER: 36,113  
 REFERENCE/DOCKET NUMBER: 600-1-142  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: 201-487-5800  
 TELEFAX: 201-343-1684  
 TELEX: 133521  
 INFORMATION FOR SEQ ID NO: 16:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 51 amino acids  
 TYPE: amino acid  
 STRANDEDNESS: single  
 TOPOLOGY: linear  
 MOLECULE TYPE: peptide  
 US-08-519-103-16

Query Match 10.3%; Score 163; DB 1; Length 51;  
 Best Local Similarity 54.9%; Pred. No. 3e-10;  
 Matches 28; Conservative 12; Mismatches 11; Indels 0; Gaps 0;

OY 65 LKRGNTEEDDELITIKLSLGNKWSLIAGRLPGRTDNEIKYNWTHIRK 115

Db 1 VKKSWTEEDRILYEAHKLGNRMALIKLPRTDINSIKHNSIMRRK 51

RESULT 11

US-08-519-103-19

Sequence 19, Application US/08519103

Patent No. 5733730

GENERAL INFORMATION:

APPLICANT: delange, Titia

TITLE OF INVENTION: TELOMERE REPEAT BINDING FACTOR AND

DIAGNOSTIC AND THERAPEUTIC USE THEREOF

NUMBER OF SEQUENCES: 21

CORRESPONDENCE ADDRESS:

ADDRESSEE: KLAUBER & JACKSON

STREET: 411 Hackensack Avenue

CITY: Hackensack

STATE: New Jersey

COUNTRY: USA

ZIP: 07601

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/519,103

FILING DATE: 25-AUG-1995

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: Crane-Feury, Sharon E.

REGISTRATION NUMBER: 36,113

REFERENCE/DOCKET NUMBER: 600-1-142

TELECOMMUNICATION INFORMATION:

TELEPHONE: 201-487-5800

TELEFAX: 201-343-1684

TELEX: 133521

INFORMATION FOR SEQ. ID NO: 19:

SEQUENCE CHARACTERISTICS:

LENGTH: 51 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: peptide

US-08-519-103-19

Query Match

Best local Similarity 10.3%; Score 163; DB 1; Length 51;

Matches 29; Conservative 12; Mismatches 10; Indels 0; Gaps 0;

Db 1 VKKSWTEEDRILYEAHKLGNRMALIKLPRTDINSIKHNSIMRRK 51

RESULT 12

US-09-018-635-16

Sequence 16, Application US/09018635

Patent No. 6297356

GENERAL INFORMATION:

APPLICANT: de Lange, Titia

APPLICANT: Broccoli, Dominique

APPLICANT: Smogorzewska, Agata

TITLE OF INVENTION: TELOMERE REPEAT BINDING FACTOR AND

DIAGNOSTIC AND THERAPEUTIC USE THEREOF

NUMBER OF SEQUENCES: 52

CORRESPONDENCE ADDRESS:

ADDRESSEE: KLAUBER & JACKSON

STREET: 411 Hackensack Avenue

CITY: Hackensack

STATE: New Jersey

COUNTRY: USA

ZIP: 07601

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/018,635

FILING DATE:

CLASSIFICATION:

ATTORNEY/AGENT INFORMATION:

NAME: David A. Jackson

REGISTRATION NUMBER: 26,742

REFERENCE/DOCKET NUMBER: 600-1-142 CIP1

TELECOMMUNICATION INFORMATION:

TELEPHONE: 201-487-5800

TELEFAX: 201-343-1684

TELEX: 133521

INFORMATION FOR SEQ. ID NO: 19:

ZIP: 07601

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/018,635

FILING DATE:

CLASSIFICATION:

ATTORNEY/AGENT INFORMATION:

NAME: David A. Jackson

REGISTRATION NUMBER: 26,742

REFERENCE/DOCKET NUMBER: 600-1-142 CIP1

TELECOMMUNICATION INFORMATION:

TELEPHONE: 201-487-5800

TELEFAX: 201-343-1684

TELEX: 133521

INFORMATION FOR SEQ. ID NO: 16:

SEQUENCE CHARACTERISTICS:

LENGTH: 51 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: peptide

US-09-018-635-16

Query Match

Best local Similarity 10.3%; Score 163; DB 4; Length 51;

Matches 28; Conservative 12; Mismatches 11; Indels 0; Gaps 0;

Db 1 VKKSWTEEDRILYEAHKLGNRMALIKLPRTDINSIKHNSIMRRK 51

RESULT 13

US-09-018-635-19

Sequence 19, Application US/09018635

Patent No. 6297356

GENERAL INFORMATION:

APPLICANT: de Lange, Titia

APPLICANT: Broccoli, Dominique

APPLICANT: Smogorzewska, Agata

TITLE OF INVENTION: TELOMERE REPEAT BINDING FACTOR AND

DIAGNOSTIC AND THERAPEUTIC USE THEREOF

NUMBER OF SEQUENCES: 52

CORRESPONDENCE ADDRESS:

ADDRESSEE: KLAUBER & JACKSON

STREET: 411 Hackensack Avenue

CITY: Hackensack

STATE: New Jersey

COUNTRY: USA

ZIP: 07601

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/018,635

FILING DATE:

CLASSIFICATION:

ATTORNEY/AGENT INFORMATION:

NAME: David A. Jackson

REGISTRATION NUMBER: 26,742

REFERENCE/DOCKET NUMBER: 600-1-142 CIP1

TELECOMMUNICATION INFORMATION:

TELEPHONE: 201-487-5800

TELEFAX: 201-343-1684

TELEX: 133521

INFORMATION FOR SEQ. ID NO: 19:

```
Search completed: January 23, 2002, 21:12:49
Job time: 3017 sec
```



GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 23, 2002, 20:25:27 ; Search time 60.99 Seconds  
(Without alignments) 367.197 Million cell updates/sec

Title: US-09-453-387A-2

Sequence: 1 MGRSPCCCEKAHTNKGAATKE.....RMCQNSNSNGFYSCRPIDS 294

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 219241 seqs, 76174552 residues

Total number of hits satisfying chosen parameters: 219241

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database :  
1: PIR1:\*  
2: PIR2:\*  
3: PIR3:\*  
4: PIR4:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match Length	DB ID	Description
1	1584	100.0	294	2 T09879 myb-related protei
2	854.5	53.9	274	1 J00957 myb-related protei
3	768	48.5	257	2 T00780 myb-related protei
4	762	48.1	226	2 D85086 DNA-binding protei
5	756.5	47.8	251	1 S35729 myb-related protei
6	728	46.0	273	1 S69189 myb-related protei
7	723	45.6	282	2 T05690 myb-related transc
8	717	45.3	282	2 T51632 myb-related transc
9	704.5	44.5	232	1 J00960 myb-related protei
10	702	44.3	246	2 T51631 probable transcrip
11	693	43.8	235	1 S04899 myb-related protei
12	684	43.2	274	2 T05769 myb-related protei
13	683	43.1	239	2 T02984 myb-related protei
14	683	43.0	267	1 S51506 myb-related protei
15	681.5	43.0	269	1 S58292 probable MYB fami
16	652	41.2	276	2 T02985 myb-related protei
17	626.5	39.6	226	2 T09773 myb-related protei
18	584.5	36.9	332	2 D86476 protein F1504.43
19	567	35.8	325	2 T51509 probable transcrip
20	549.5	34.7	282	2 T85327 myb-related protei
21	546.5	34.5	274	2 T07393 myb-related prote
22	543	34.3	414	2 A86229 hypothetical prote
23	537	33.9	368	2 T03828 myb protein - rice
24	534.5	33.7	324	2 B85064 MYB-like protein
25	533.5	33.7	326	2 T49966 myb-related protei
26	530	33.5	282	2 C96687 hypothetical protei
27	528.5	33.4	352	1 S58293 myb-related protei
28	528	33.3	421	1 S36605 myb-related protei
29	526	33.2	314	2 E96609 probable DNA-bind

30	524	33.1	249	1 S68688 myb-related protei
31	522	33.0	299	2 T47917 probable transcrip
32	520.5	32.9	369	2 T01196 transcription fact
33	518.5	32.7	336	2 T51644 probable transcrip
34	517.5	32.7	352	2 T51659 myb-related transc
35	515.5	32.5	334	2 T50816 probable transcrip
36	515.5	32.5	371	2 T00438 probable MYB fami
37	514	32.4	453	2 T09745 myb-related protei
38	513	32.4	333	2 G96768 protein transcrip
39	511.5	32.3	371	2 T51636 myb-related transc
40	510	32.2	365	2 D86470 hypothetical prote
41	509	32.1	278	2 T51641 myb-related transc
42	509	32.1	421	1 S24244 myb-related protei
43	507.5	32.0	249	2 E84717 probable MYB fami
44	507.5	32.0	263	2 T48607 probable transcrip
45	505.5	31.9	343	2 T48050 probable transcrip

## ALIGNMENTS

RESULT 1.  
T09879  
myb-related protein A - upland cotton  
C:Species: Gossypium hirsutum (upland cotton)  
C>Date: 16-Jul-1999 #sequence\_revision 16-Jul-1999 #text\_change 26-Aug-1999  
C:Accession: T09879  
R:Wilkins, T.A.; Lu, C.C.  
Submitted to the EMBL Data Library, January 1994  
A:Reference number: Z16894  
A:Accession: T09879  
A>Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: mRNA  
A:Residues: 1-294 <WLD>  
A:Cross-references: EMBL: L04497; NID: g437326; PID: g437327  
C:Genetics:  
A:Gene: MYB A  
C:Superfamily: barley myb-related protein 1; myb DNA-binding repeat homology  
C:Keywords: DNA binding; transcription regulation  
F:62-112/Domain: myb DNA-binding repeat homology <MYB>

Query Match	100.0%	Score 1584	DB 2	Length 294
Best Local Similarity	100.0%	Pred No. 2e-111		
Matches 294	Conservative	0	Mismatches 0	Indels 0
Gaps 0				
QY	1	MGRSPCCCEKAHTNKGAATKEEDQRLINIRVHGGCWRSLPKAAGLRGCKSRLRWINY	60	
DB	1	MGRSPCCCEKAHTNKGAATKEEDQRLINIRVHGGCWRSLPKAAGLRGCKSRLRWINY	60	
QY	61	LRPDIKRGNTFEDELLIKLHSLGNKWSLIGRLPGRDNEIKYWNTHIKRKLISRG	120	
DB	61	LRPDIKRGNTFEDELLIKLHSLGNKWSLIGRLPGRDNEIKYWNTHIKRKLISRG	120	
QY	121	IDPQHRPLNQTANTVTWAPFELDFRNSPTYSKSSSINKPSLDFVYNEHOFKSNPDSL	180	
DB	121	IDPQHRPLNQTANTVTWAPFELDFRNSPTYSKSSSINKPSLDFVYNEHOFKSNPDSL	180	
QY	181	EEPNCTASSGNTTDEEQDQLHKKQYGPNSGQDINELISGIYSADSRVSNANSASEK	240	
DB	181	EEPNCTASSGNTTDEEQDQLHKKQYGPNSGQDINELISGIYSADSRVSNANSASEK	240	
QY	241	PRVDNNNFQLEQAMVAKAVCLQWJGFGTSEICRNCQNSNSNGFYSCRPIDS	294	
DB	241	PRVDNNNFQLEQAMVAKAVCLQWJGFGTSEICRNCQNSNSNGFYSCRPIDS	294	
RESULT 2				
J00957				
myb-related protein 330 - garden snapdragon				
C:Species: Antirrhinum majus (garden snapdragon)				
C>Date: 30-Sep-1991 #sequence_revision 30-Sep-1991 #text_change 24-Jul-1997				
C:Accession: J00957				

R.Jackson, D.; Cullanez-Macia, F.; Prescott, A.G.; Roberts, K.; Martin, C.  
Plant Cell 3, 115-125, 1991  
A:Title: Expression patterns of myb genes from Antirrhinum flowers.  
A:Reference number: J00956; MUID:93005689  
A:Accession: J00957

A:Molecule type: mRNA  
A:Residues: 1-274 <JAC>  
A:Experimental source: strain J1:522  
C:Superfamily: barley myb-related protein 1; myb DNA-binding repeat homology  
C:Keywords: DNA binding; duplication; nucleus; transcription regulation  
F:9-61/Domain: myb DNA-binding repeat homology <MYB1>  
F:62-112/Domain: myb DNA-binding repeat homology <MYB2>

Query Match 53.9%; Score 854.5; DB 1; Length 274;  
Best Local Similarity 60.5%; Pred. No. 7.6e-57;  
Matches 178; Conservative 23; Mismatches 68; Indels 25; Gaps 8;

OY 1 MGRSPCEKAHTNKGAWTKEEDQRLINIRVHGGCWSLPPKAGLLRCGKSCRLRWNY 60  
DB 1 MGRSPCEKAHTNKGAWTKEEDQRLINIRVHGGCWSLPPKAGLLRCGKSCRLRWNY 60  
OY 61 LRPDLKRGNTFEEDDELIIKLSLGNKWSLIGRLPGRTDNEIKYNTWHIKRKLISRG 120  
DB 61 LRPDLKRGNTFEEDDELIIKLSLGNKWSLIGRLPGRTDNEIKYNTWHIKRKLISRG 120  
OY 121 IDPQTHRLNQTANTNTVTAPE---LDFRNSPTSVKSSSIKNSLDFENEFQKSN 176  
DB 121 IDPQTHRLNQTANTNTVTAPE---LDFRNSPTSVKSSSIKNSLDFENEFQKSN 176  
OY 177 TDSLEPNCATSSGWTDEEQEQLHKQOYGPSNGODINLELSIGIVASDSRVSNS 236  
DB 171 SSSSTDPTCKNSS---TTEESQSLITPPKREKSVPLVDLSLGLPQSOCKNSVSLN 226  
OY 237 AESKPKVDNNFQLEQAMAKAVCLC-WOLGFGTSEICRNQNSNGFYSC 289  
DB 227 SSSSGFYD---FRRPRAKVAQRMCVCWTLGLQKGEQFCNQ--SFNGFYRYC 274

## RESULT 3

myb-related protein T22J18.19 - Arabidopsis thaliana  
C:Species: Arabidopsis thaliana (mouse-ear cress)  
C>Date: 12-Feb-1999 #sequence\_revision 12-Feb-1999 #text\_change 22-Oct-1999  
A:Accession: T00780  
R:Vignatara, V.S.; Schwartz, J.R.; Tortum, M.; Yu, G.; Oji, O.; Kwan, A.; Liu, S.; Li, Tz, D.; Li, Y.; Palm, C.J.; Shinn, P.; Sun, H.; Davis, R.W.; Eckert, J.R.; Federspiel, R.  
submitted to the EMBL Data Library, July 1998  
A:Description: Arabidopsis thaliana chromosome 1 BAC T22J18 sequence, complete sequence.  
A:Reference number: 214202  
A:Accession: T00780  
A:Status: translated from GB/EMBL/DBD  
A:Molecule type: DNA  
A:Residues: 1-257 <YVS>  
A:Cross-references: EMBL:AC003979; NID:g3172156; PIDN:AAC25522.1; PID:g32694; GSPDB:GN00  
A:Experimental source: cultivar Columbia  
A:Genetics:  
A:Gene: AT2G218.19  
A:Map position: 1  
A:Initons: 45/1; 88/2  
C:Superfamily: barley myb-related protein 1; myb DNA-binding repeat homology  
C:Keywords: DNA binding; duplication; nucleus; transcription regulation  
F:9-61/Domain: myb DNA-binding repeat homology <MYB1>  
F:62-112/Domain: myb DNA-binding repeat homology <MYB2>

Query Match 48.5%; Score 768; DB 2; Length 257;  
Best Local Similarity 55.6%; Pred. No. 2.1e-50;  
Matches 160; Conservative 33; Mismatches 53; Indels 42; Gaps 7;

OY 1 MGRSPCEKAHTNKGAWTKEEDQRLINIRVHGGCWSLPPKAGLLRCGKSCRLRWNY 60  
DB 1 MGRSPCEKAHTNKGAWTKEEDQRLINIRVHGGCWSLPPKAGLLRCGKSCRLRWNY 60

OY 61 LRPDLKRGNTFEEDDELIIKLSLGNKWSLIGRLPGRTDNEIKYNTWHIKRKLISRG 120  
DB 61 LRPDLKRGNTFEEDDELIIKLSLGNKWSLIGRLPGRTDNEIKYNTWHIKRKLISRG 120  
OY 121 IDPQTHRLNQTANTNTVTAPELDFRNSPTSVKSSSIKNSLDFENEFQKSN 180  
DB 121 IDPQTHRLNQTANTNTVTAPELDFRNSPTSVKSSSIKNSLDFENEFQKSN 180  
OY 181 EEP-----NCTASSGWTDEEQEQLHKQOYGPSNGODINLELSIG-IVSADSSR-VSN 233  
DB 161 REELGMVNNC-ESSSTTSKD-----YNEEDWVNLIELSVSPSTRESTRKYSV 209  
OY 234 ANSAEKPKVDNNFQLEQAMAKAVCLC-WOLGFGTSEICRNQNSN 281  
DB 210 VDSAESTRMGESEF---GAHESDAVCCRCRGLPNEESCRCNRVSD 253

## RESULT 4

D85096  
DNA-binding protein [imported] - Arabidopsis thaliana  
C:Species: Arabidopsis thaliana (mouse-ear cress)  
C>Date: 16-Feb-2001 #sequence\_revision 16-Feb-2001 #text\_change 02-Mar-2001  
A:Accession: D85096  
R:anonymous, The European Union Arabidopsis Genome Sequencing Consortium, The Cold Nature 402, 769-777, 1999  
A:Title: Sequence and analysis of chromosome 4 of the plant Arabidopsis thaliana.  
A:Reference number: A85001; MUID:20083488  
A:Accession: D85096  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-236 <STO>  
A:Cross-references: GB:NC\_001268; NID:g7267641; PIDN:CA878069.1; GSPDB:GN00140  
A:Genetics:  
A:Gene: AT4G09460  
A:Map position: 4  
C:Superfamily: Arabidopsis myb-related protein 5; myb DNA-binding repeat homology

Query Match 48.1%; Score 762; DB 2; Length 236;  
Best Local Similarity 58.9%; Pred. No. 5.3e-50;  
Matches 145; Conservative 39; Mismatches 40; Indels 22; Gaps 5;

OY 1 MGRSPCEKAHTNKGAWTKEEDQRLINIRVHGGCWSLPPKAGLLRCGKSCRLRWNY 60  
DB 1 MGRSPCEKAHTNKGAWTKEEDQRLINIRVHGGCWSLPPKAGLLRCGKSCRLRWNY 60  
OY 61 LRPDLKRGNTFEEDDELIIKLSLGNKWSLIGRLPGRTDNEIKYNTWHIKRKLISRG 120  
DB 61 LRPDLKRGNTFEEDDELIIKLSLGNKWSLIGRLPGRTDNEIKYNTWHIKRKLISRG 120  
OY 121 IDPQTHRLNQTANTNTVTAPELDFRNSPTSVKSSSIKNSLDFENEFQKSN 179  
DB 121 IDPQTHRLNQTANTNTVTAPELDFRNSPTSVKSSSIKNSLDFENEFQKSN 179  
OY 180 LEEPNCATSSGWTDEEQEQLHKQOYGPSNGO-DINLELSIGIVASDSRVSNSAE 238  
DB 164 SSD-NGASTSGTTDEDLRQ--NGECYSDNSGHINKINDLLTGLFGWSMGRIVGVSSAD 220  
OY 239 SKPKVD 244  
DB 221 SKPMCD 226

## RESULT 5

S35729  
myb-related protein 2 - barley  
N:Alternate names: myb2 protein; MybH5 protein  
C:Species: Hordeum vulgare (Barley)  
C>Date: 03-Mar-1994 #sequence\_revision 26-May-1995 #text\_change 18-Jun-1999  
A:Accession: S35729; S31813; S61507  
R:Roide, W.  
submitted to the EMBL Data Library, January 1993

A:Reference number: S35419  
 A:Accession: S35729  
 A:Molecule type: mRNA  
 A:Residues: 1-251 <ROH>  
 A:Cross-references: EMBL:X70876; NID:g19054; PIDN:CAA50221.1; PID:g19055  
 R:Wissenbach, M.; Ueberlacker, B.; Vogt, F.; Backer, D.; Salami, F.; Rohde, W.  
 submitted to the EMBL Data Library, January 1993  
 A:Description: Myb gene from hordeum vulgare: tissue-specific expression of chimeric Myb  
 A:Reference number: S31813  
 A:Accession: S31813  
 A:Molecule type: DNA  
 A:Residues: 1-104 <MIS>  
 A:Cross-references: EMBL:X70880; NID:g19056; PIDN:CAA50225.1; PID:g19057  
 R:Wissenbach, M.; Ueberlacker, B.; Vogt, F.; Becker, D.; Salami, F.; Rohde, W.  
 Plant J. 4, 411-422, 1993  
 A:Title: Myb genes from Hordeum vulgare: tissue-specific expression of chimeric Myb prot  
 A:Reference number: S61507  
 A:Accession: S61507  
 A:Status: preliminary; nucleic acid sequence not shown  
 A:Molecule type: DNA  
 A:Residues: 1-117 <M12>  
 A:Cross-references: EMBL:X70880  
 C:Genetics:  
 A:Gene: myb2  
 A:Introns: 88/2  
 C:Superfamily: barley myb-related protein 1; myb DNA-binding repeat homology  
 C:Keywords: DNA binding; duplication; nucleus; transcription regulation  
 F:9-61/Domain: myb DNA-binding repeat homology <MYB>  
 F:62-112/Domain: myb DNA-binding repeat homology <MYB2>

Query Match 47.8%; Score 756.5; DB 1; Length 251;  
 Best Local Similarity 50.7%; Pred. No. 1.5e-49;

Matches 152; Conservative 32; Mismatches 55; Indels 61; Gaps 6;

OY 1 MGRSPCCERKATNKGAATKEEDORLINTYRVHGGCNRSLPKAAGLLRCGKSCRLRWNY 60  
 |||||  
 DB 1 MGRSPCCERKATNKGAATKEEDORLINTYRVHGGCNRSLPKAAGLLRCGKSCRLRWNY 60  
 OY 61 LRPDLKRGNTFEEDDELLIKLHSLGNKWSLJAGRLPGRTDNEIKNYWTHIRKRLISRG 120  
 |||||  
 DB 61 LRPDLKRGNTFEEDDELLIKLHSLGNKWSLJAGRLPGRTDNEIKNYWTHIRKRLISRG 120  
 OY 121 IDPQTHRLPLNOTANTNTVTAPEL--DFRNSPTSV--SKSSSIKNPSLDYFNNEFOKSNST 177  
 |||||  
 DB 121 IDPQTHRLPLNOTANTNTVTAPEL--DFRNSPTSV--SKSSSIKNPSLDYFNNEFOKSNST 177  
 OY 178 DSLEENCTASSGWTDEEQOQLHKQOYGPSNGODINLELSIGIVSADSRVSNANSA 237  
 |||||  
 DB 181 TSTGGERC-----PDNLNLDLSVGPADTP----- 205  
 OY 238 ESKPKYDNNNFOLFEOAMVAKAVLCWOLGFGTSEICRNCQNSNS---NGFSTYCRPLD 293  
 |||||  
 DB 206 -----TSHPVCLCRHLGFRGEGAC--SCROADSAGSOGGGRFYRFRPLE 246

## RESULT 6

S69189  
 myb-related protein TMH27 - tomato  
 C:Species: Lycopersicon esculentum (tomato)  
 C:Date: 06-Dec-1996 #sequence\_revision 13-Mar-1997 #text\_change 16-Jun-2000  
 R:Lin, Q.; Hamilton, W.D.O.; Merryweather, A.  
 Plant Mol. Biol. 30, 1009-1020, 1996  
 A:Title: Cloning and initial characterization of 14 myb-related cDNAs from tomato (Lycop  
 A:Reference number: S69189; MUID:96270378  
 A:Accession: S69189  
 A:Status: preliminary  
 A:Molecule type: mRNA  
 A:Residues: 1-273 <LIN>  
 A:Cross-references: EMBL:X95296; NID:g1167483; PIDN:CAA64614.1; PID:g1167484  
 C:Genetics:  
 A:Gene: TMH27

C:Superfamily: barley myb-related protein 1; myb DNA-binding repeat homology  
 C:Keywords: DNA binding; duplication; nucleus; transcription regulation  
 F:9-61/Domain: myb DNA-binding repeat homology <MYB>  
 F:62-112/Domain: myb DNA-binding repeat homology <MYB2>

Query Match 46.0%; Score 728; DB 1; Length 273;  
 Best Local Similarity 53.0%; Pred. No. 2.2e-47;  
 Matches 151; Conservative 26; Mismatches 64; Indels 44; Gaps 5;

OY 1 MGRSPCCERKATNKGAATKEEDORLINTYRVHGGCNRSLPKAAGLLRCGKSCRLRWNY 60  
 |||||  
 DB 1 MGRSPCCERKATNKGAATKEEDORLINTYRVHGGCNRSLPKAAGLLRCGKSCRLRWNY 60  
 OY 61 LRPDLKRGNTFEEDDELLIKLHSLGNKWSLJAGRLPGRTDNEIKNYWTHIRKRLISRG 120  
 |||||  
 DB 61 LRPDLKRGNTFEEDDELLIKLHSLGNKWSLJAGRLPGRTDNEIKNYWTHIRKRLISRG 120  
 OY 121 IDPQTHRLPLNOTANTNTVTAPELDFRNSPTSVSKSSSIKNPSLDYFNNEFOKSNSTDL 180  
 |||||  
 DB 121 IDPQTHRLPLNOTANTNTVTAPELDFRNSPTSVSKSSSIKNPSLDYFNNEFOKSNSTDL 180  
 OY 181 EEPNCTASSGWTDEEQOQLHKQOYGPSNGODINLELSIGIVSADSRVSNANSAESK 240  
 |||||  
 DB 158 EMINIKAEFVETSKESDNEI--IQKSSSCLDPLNLELTI-----SPPH 201  
 OY 241 PKVDNNNFOLFEOAMVAKAVLCWOLGFGTSEICRNCQNSNSNGF 285  
 |||||  
 DB 202 QQLDHRHH-----QRSSLCFTCSIGIQNSKCSGSGSENGNW 241

## RESULT 7

T05690  
 myb-related transcription factor MYB4 - Arabidopsis thaliana  
 N:Alternate names: protein F20M13.180  
 C:Species: Arabidopsis thaliana (mouse-ear cress)  
 C:Date: 23-Apr-1999 #sequence\_revision 23-Apr-1999 #text\_change 13-Aug-1999  
 R:Bevan, M.; Wedler, H.; Kutzner, M.; Wambutt, R.; Bancroft, I.; Mewes, H.W.; Mayer, submitted to the Protein Sequence Database, February 1999  
 A:Reference number: Z15420  
 A:Accession: T05690  
 A:Molecule type: DNA  
 A:Residues: 1-282 <BEV>  
 A:Cross-references: EMBL:AL035540  
 C:Experimental source: cultivar Columbia; BAC clone F20M13  
 A:Gene: MYB4  
 A:Map position: 4  
 A:Introns: 88/2  
 A:Note: F20M13.180  
 C:Superfamily: barley myb-related protein 1; myb DNA-binding repeat homology  
 C:Keywords: DNA binding; duplication; nucleus; transcription regulation  
 F:9-61/Domain: myb DNA-binding repeat homology <MYB1>  
 F:62-112/Domain: myb DNA-binding repeat homology <MYB2>

Query Match 45.6%; Score 723; DB 2; Length 282;  
 Best Local Similarity 52.0%; Pred. No. 5.5e-47;

Matches 146; Conservative 38; Mismatches 49; Indels 48; Gaps 7;

OY 1 MGRSPCCERKATNKGAATKEEDORLINTYRVHGGCNRSLPKAAGLLRCGKSCRLRWNY 60  
 |||||  
 DB 1 MGRSPCCERKATNKGAATKEEDORLINTYRVHGGCNRSLPKAAGLLRCGKSCRLRWNY 60  
 OY 61 LRPDLKRGNTFEEDDELLIKLHSLGNKWSLJAGRLPGRTDNEIKNYWTHIRKRLISRG 120  
 |||||  
 DB 61 LRPDLKRGNTFEEDDELLIKLHSLGNKWSLJAGRLPGRTDNEIKNYWTHIRKRLISRG 120  
 OY 121 IDPQTHRLPLNOTANTNTVTAPELDFRNSPT--SVKSSSIKNPSLDYFNNEFOKSNSTDS 179  
 |||||  
 DB 121 IDPQTHRLPLNOTANTNTVTAPELDFRNSPT--SVKSSSIKNPSLDYFNNEFOKSNSTDS 176

OY 180 LEEPNCATSSGTTDEEQEOLHKKQOYGPSNGODINLEL----- 219  
 DB 177 I-----SMUTFKREKDECPVQEK-----PDLNLELISLPDDVRLQGHGKSTTP 222  
 OY 220 -----SIGIVSADSSRVSNAN-----SAESKPKVNNNFOFL 251  
 DB 223 RCFKCSLGMINGMECHRCGRMRCDDVGVSSKSGSDMSNGDFL 263

## RESULT 8

TS1632  
 myb-related transcription factor MYB4 [imported] - Arabidopsis thaliana  
 C:Species: Arabidopsis thaliana (mouse-ear cress)  
 C:Date: 18-Aug-2000 #sequence\_rev18-Aug-2000 #text\_change 02-Sep-2000  
 C:Accession: T51632  
 R:Kranz, H.D.; Denekamp, M.; Greco, R.; Jin, H.; Kranz, H.D.; Denekamp, M.; Greco, R.; J  
 ; Paz-Ares, J.; Weisshaar, B.  
 Plant J. 16, 263-276, 1998  
 A:Title: Towards functional characterisation of the members of the R2R3-MYB gene from Ar  
 A:Reference number: Z14349; MUID:9839469  
 A:Accession: T51632  
 A:Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: mRNA  
 A:Residues: 1-262 <KRA>  
 A:Cross-references: EMBL:AF062860; PIDN:AAC83582.1  
 A:Experimental source: cultivar Columbia  
 C:Genetics:  
 A:Gene: MYB4  
 C:Superfamily: barley myb-related protein 1; myb DNA-binding repeat homology

Query Match 45.3%; Score 717; DB 2; Length 282;  
 Best Local Similarity 51.6%; Pred. No. 1.6e-46;  
 Matches 145; Conservative 38; Mismatches 50; Indels 48; Gaps 7;

OY 1 MGRSPCCERKATNKGAATKEDORLNIYRVHGECCWRSPLKAGLLRCGKSCRLRWINY 60  
 DB 1 MGRSPCCERKATNKGAATKEDORLNIYRVHGECCWRSPLKAGLLRCGKSCRLRWINY 60  
 OY 61 LRPDLKRGNTFEEDELLIKHLSLGNKWSLIAGRLPGRTDNEIKNYNTHIRKRLISRG 120  
 DB 61 LRPDLKRGNTFEEDELLIKHLSLGNKWSLIAGRLPGRTDNEIKNYNTHIRKRLISRG 120  
 OY 121 IDPOTHRPLNO-TANTNTVTAFTLDFRNSPT-SYKSSSINKPSLDFVNEFOFKSNTDS 179  
 DB 121 IDPOTHRPLNO-TANTNTVTAFTLDFRNSPT-SYKSSSINKPSLDFVNEFOFKSNTDS 179  
 OY 121 IDPOTHRPLNO-TANTNTVTAFTLDFRNSPT-SYKSSSINKPSLDFVNEFOFKSNTDS 179  
 DB 121 IDPOTHRPLNO-TANTNTVTAFTLDFRNSPT-SYKSSSINKPSLDFVNEFOFKSNTDS 179  
 OY 180 LEEPNCATSSGTTDEEQEOLHKKQOYGPSNGODINLEL----- 219  
 DB 177 I-----SMUTFKREKDECPVQEK-----PDLNLELISLPDDVRLQGHGKSTTP 222  
 OY 220 -----SIGIVSADSSRVSNAN-----SAESKPKVNNNFOFL 251  
 DB 223 RCFKCSLGMINGMECHRCGRMRCDDVGVSSKSGSDMSNGDFL 263

## RESULT 9

J00960  
 myb-related protein 308 - garden snapdragon  
 C:Species: Antirrhinum majus (garden snapdragon)  
 C:Date: 30-Sep-1991 #sequence\_rev18-Aug-1991 #text\_change 24-Jul-1997  
 C:Accession: J00960  
 R:Jackson, D.; Cullane-Macla, F.; Prescott, A.G.; Roberts, K.; Martin, C.  
 Plant Cell 3, 115-125, 1991  
 A:Title: Expression patterns of myb genes from Antirrhinum flowers.  
 A:Reference number: J00956; MUID:93005689  
 A:Accession: J00960  
 A:Molecule type: mRNA  
 A:Residues: 1-232 <JAC>  
 A:Experimental source: strain J1-522  
 C:Comment: The gene encoding for this protein is expressed in all plant organs.  
 C:Superfamily: barley myb-related protein 1; myb DNA-binding repeat homology  
 C:Keywords: DNA binding; duplication; nucleus; transcription regulation

F:9-61/Domain: myb DNA-binding repeat homology <MYB1>  
 F:62-112/Domain: myb DNA-binding repeat homology <MYB2>

Query Match 44.5%; Score 704.5; DB 1; Length 232;  
 Best Local Similarity 52.4%; Pred. No. 1.1e-45;  
 Matches 150; Conservative 24; Mismatches 55; Indels 57; Gaps 8;

OY 1 MGRSPCCERKATNKGAATKEDORLNIYRVHGECCWRSPLKAGLLRCGKSCRLRWINY 60  
 DB 1 MGRSPCCERKATNKGAATKEDORLNIYRVHGECCWRSPLKAGLLRCGKSCRLRWINY 60  
 OY 61 LRPDLKRGNTFEEDELLIKHLSLGNKWSLIAGRLPGRTDNEIKNYNTHIRKRLISRG 120  
 DB 61 LRPDLKRGNTFEEDELLIKHLSLGNKWSLIAGRLPGRTDNEIKNYNTHIRKRLISRG 120  
 OY 121 IDPOTHRPLNO-TANTNTVTAFTLDFRNSPT-SYKSSSINKPSLDFVNEFOFKSNTDS 179  
 DB 121 IDPOTHRPLNO-TANTNTVTAFTLDFRNSPT-SYKSSSINKPSLDFVNEFOFKSNTDS 179  
 OY 180 LEEPNCATSSGTTDEEQEOLHKKQOYGPSNGO--DINLELSIGIVSADSSRVSNANSA 237  
 DB 150 KEE-----DTKHKAVADIMIKENSPVQERCPDNLNLT-----SPCCQOQINTHOE 197  
 OY 238 ESKPKVNNNFOFLQAVAKAVCLQWLGFTSEICRNCQNSN 283  
 DB 198 NLKTGGRNG-----SPTLCFVCRLGIONSKDC-SCSDGVN 232

## RESULT 10

TS1631  
 probable transcription factor MYB3 [imported] - Arabidopsis thaliana  
 C:Species: Arabidopsis thaliana (mouse-ear cress)  
 C:Date: 18-Aug-2000 #sequence\_rev18-Aug-2000 #text\_change 02-Sep-2000  
 C:Accession: T51631  
 R:Kranz, H.D.; Denekamp, M.; Greco, R.; Jin, H.; Kranz, H.D.; Denekamp, M.; Greco, R.  
 ; Paz-Ares, J.; Weisshaar, B.  
 Plant J. 16, 263-276, 1998  
 A:Title: Towards functional characterisation of the members of the R2R3-MYB gene fr  
 A:Reference number: Z14349; MUID:9839469  
 A:Accession: T51631  
 A:Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: mRNA  
 A:Residues: 1-246 <KRA>  
 A:Cross-references: EMBL:AF062859; PIDN:AAC83581.1  
 A:Experimental source: cultivar Columbia  
 C:Genetics:  
 A:Gene: MYB3  
 C:Superfamily: barley myb-related protein 1; myb DNA-binding repeat homology

Query Match 44.3%; Score 702; DB 2; Length 246;  
 Best Local Similarity 54.0%; Pred. No. 1.7e-45;  
 Matches 149; Conservative 33; Mismatches 52; Indels 42; Gaps 7;

OY 13 NKGAATKEDORLNIYRVHGECCWRSPLKAGLLRCGKSCRLRWINYLRPDLKRGNTF 72  
 DB 2 NKGAATKEDORLNIYRVHGECCWRSPLKAGLLRCGKSCRLRWINYLRPDLKRGNTF 61  
 OY 73 EDELLIKHLSLGNKWSLIAGRLPGRTDNEIKNYNTHIRKRLISGIDPQTHRLNQT 132  
 DB 62 EDELLIKHLSLGNKWSLIAGRLPGRTDNEIKNYNTHIRKRLISGIDPQTHRLNQT 120  
 OY 133 ANTNTVTAFTLDFRNSPT-SYKSSSINKPSLDFVNEFOFKSNTDSLEP-----NCTA 187  
 DB 121 -----SVSPSSLQWDVETLHDSGPVAPPEVREIGAVNAC-E 160  
 OY 188 SSGMTTDEEQEOLHKKQOYGPSNGODINLELSIG-IVSADSSR-VSNANSAESKPKVDN 245  
 DB 161 SSGTSEKD-----YGNEDMVNLDELISVGSYKRESTRKYSVDSABSTRMGS 210  
 OY 246 NNFOFLQAVAKAVCLQWLGFTSEICRNCQNSN 281



A:Reference number: S61506; MUID:94035190  
A:Accession: S61506  
A:Status: preliminary.  
A:Molecule type: DNA  
A:Residues: 1-267 <MIS>  
A:Cross-references: EMBL:X70879; NID:g19052; PIDN:CAA50224.1; PID:g19053  
A:Experimental source: var. Abyssinian 2231  
R:Marocco, A.; Wissenbach, M.; Becker, D.; Paz-Ares, J.; Saedler, H.; Salamini, F.; Rohde  
Mol. Gen. Genet. 216, 183-187, 1989  
A:Title: Multiple genes are transcribed in Hordeum vulgare and Zea mays that carry the D  
A:Reference number: S04896; MUID:89313655  
A:Accession: S04896  
A:Molecule type: mRNA  
A:Residues: 1-267 <MAR>  
A:Cross-references: EMBL:X70877; NID:g19050; PIDN:CAA50222.1; PID:g19051  
A:Experimental source: var. Abyssinian 2231  
C:Genetics:  
A:Introns: 88/2  
C:Superfamily: barley myb-related protein 1; myb DNA-binding repeat homology  
C:Keywords: DNA binding; duplication; nucleus; transcription regulation  
F:9-61/Domain: myb DNA-binding repeat homology <MYB1>  
F:62-112/Domain: myb DNA-binding repeat homology <MYB2>

```

Query Match Similarity      43.1%; Score 683; DB 1; Length 267;
Best Local Identity         50.7%; Pred. NO. 5.1e-44;
Matches 142; Conservative  28; Mismatches 72; Indels 38; Gaps 5

Qy      1 MGRSPCCCEKAHFNKAMWKEEDQRLINTYRVAGEGCMRSLPKAAGLLRCGKSCRLMWNY 60
      |||||
Db      1 MGRSPCCCEKAHFNKAMWKEEDRLTATYKANGEGCMRSLPKAAGLLRCGKSCRLMWNY 60
      |||||

Qy      61 LRPDLKRGNFTEEBDELIIKLHSLGKNKWSL IAGRLPGRTDNEIKNYWTHIRKLLISG 120
      |||||
Db      61 LRPDLKRGNF SHEEDELIIKLHSLGKNKWSL IAGRLPGRTDNEIKNYWTHIRKLLISG 120
      |||||

Qy      121 IDPQTHRLINOTANTNTWTATPTLEDFRNSPTYSKVSSTIKNPSLDENYNEPQKSTDSI 180
      |||||
Db      121 IDPVTHTRAINSDDHAASNTISFESAQRDDKGA VFR-----RDAAE 159
      |||||

Qy      161 EEPNCTASSGMTTDEEQDQBLHKKQOYG---PSNGODINLELSIGIVSADSSRYSANNSA 237
      |||||
Db      160 PAKAAAAAAAIHHVDDHHHRSNPOLDWVGKRLKCPDNLMDICI-----SPEDHPDV 213
      |||||

Qy      238 ESKPKYDNNNFQFLBDQAVAKAYCLCW---QLGPGTSEIOR 275
      |||||
Db      214 DTKPYVKKR-----EAGVGVGVAAGLDFSCSMGLPNSDDCK 247
      |||||

```

RESULT 15  
 558292  
 probable MYB family transcription factor At2g16720 [imported] - Arabidopsis thaliana  
 N:Alternate names: myb-related protein 149  
 C:Species: Arabidopsis thaliana (mouse-ear cress)  
 C:Date: 13-Jan-1996 #sequence\_revision 19-Apr-1996 #text\_change 02-Mar-2001  
 C:Accession: S58292; D84543  
 R:Quenedvliet, N.E.M.; Dockx, J.; Keultjes, G.G.M.; Smeekens, J.C.M.  
 submitted to the EMBL Data Library, July 1995  
 A:Reference number: S58280  
 A:Accession: S58292  
 A:Molecule type: DNA  
 A:Residues: 1-269 <DNA>  
 A:Cross-references: EMBL:X90385; NID:G928968; PIDD:CAA62033.1; PIDA:G928966  
 R:Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Beito, M.T.; Town, C.D.; Fujii, C.Y.;  
 W.; Koo, H.; Wolfat, K.S.; Cronin, L.A.; Shen, M.; Vanden, S.E.; Umagum, L.; Tallon,  
 M.; Doo, J.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter,  
 Nature 402, 761-768, 1999  
 A:Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.  
 A:Reference number: A84420; MUID:20083487  
 A:Accession: D84543  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-269 <STO>

A:Cross-references: GB:AE002093; NID:g4581115; PIDN:AAD24605.1; GSPDB:GN001333  
 C:Genetics:  
 A:Gene: At2g16720  
 A:Map position: 2  
 A:Introns: 88/2  
 C:Superfamily: barley myb-related protein 1; myb DNA-binding repeat homology  
 C:Keywords: DNA binding; duplication; nucleus; transcription regulation  
 E:67-101/Domain: myb DNA-binding repeat homology <MYB1>  
 E:67-112/Domain: myb DNA-binding repeat homology <MYB2>

	Query Match	43.0%	Score 681.5	DB 1.	Length 269;	
	Best Local Similarity	48.8%	Pred. No. 6.7e-44;			
	Matches 142;	Conservative 35;	Mismatches 59;	Indels 55;	Gaps	8
Qy	1 MGSPCCCEKATNKGAMTKEEDORLINTYIRVHGGCGWRLPKAGALLRCGSKCRLRWNTY		:::::		:::::	60
Dd	1 MGSPPCEEKEMNNKGAWTKEEDELRVSYIKSHHGCGWRSLPRAAGLLRCGSKCRLRWNTY		:::::		:::::	60
Qy	61 LRPLDKGNFTLEEBDELIIKLHSLLGNKWSLIAGRLLPGRTDNEIKYNWNTHIKRKLISRG		:::::		:::::	120
Dd	61 LRPLDKGNFNLHDEBELIILKLSLIGNKMSLIARLPGRTDNEIKYNWNTHIKRKLISKG		:::::		:::::	120
Qy	121 IDPOTHRPLNOTANTNTVTAPTEELDERNSPTYSKSSSINPSIDFNENPEQFSKNTSTSL		:::::		:::::	180
Dd	121 IDEATHGININEA-----KISDLKKTKQOIVDVSVTFKFE-----TDKSGD		:::::		:::::	162
Qy	181 EEPICCTASSGGTTTDEEQ--QEQLHKQOYGPNQGODINIELSIGIYSADSRVSANMISA		:::::		:::::	237
Dd	163 QOKNKYIRNGEYVCKEERVAVEEKI-----GPDNLIELRLSPPMQNOREISTCTAS		:::::		:::::	212
Qy	238 ESARKKVONNNPFLEOAANVAKAVCLWGGLGFSEICRNCONSNNGFYSY 288		:::::		:::::	
Dd	213 RF-----YENDMEC-----SSEYVK-CQTESSSSI-ST 239		:::::		:::::	

Search completed: January 23, 2002, 21:14:02  
Job time: 2915 sec

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 23, 2002, 21:12:53 ; Search time 51.05 Seconds  
(without alignments)  
211.155 Million cell updates/sec

File: US-09-453-387a-2

Perfect score: 1584  
Sequence: 1 MGRSPCCERAHNTKGMATKE.....RNCONSNNGFYSCRPIDS 294

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 100059 seqs, 36664827 residues  
Total number of hits satisfying chosen parameters: 100059

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : SwissProt\_39:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	693	43.8	255	1 MYB3_MAIZE	P20025 zea mays (m
2	683	43.1	267	1 MYB1_HORVU	P20026 hordeum vul
3	509	32.1	421	1 MYBP_PHYPA	P80073 physcomitre
4	496.5	31.3	399	1 MYBP_MAIZE	P27898 zea mays (m
5	490	30.9	340	1 MYB1_MAIZE	P20024 zea mays (m
6	454.5	28.0	302	1 MYB3_HORVU	P20027 hordeum vul
7	443.5	27.7	273	1 MYBC_MAIZE	P20390 zea mays (m
8	439	27.4	252	1 MYBD_MAIZE	P27900 arabidopsis
9	402.5	25.4	238	1 G11_ARATH	P01104 avian myelo
10	332	20.3	382	1 MYBB_AVIMB	P03337 gallus gall
11	312	19.7	686	1 MYBB_CHICK	P05935 xenopus lae
12	309	19.5	728	1 MYBA_XENLA	P52551 xenopus lae
13	306.5	19.3	743	1 MYBB_XENLA	P48972 zea mays (m
14	304.5	19.2	704	1 MYBB_MOUSE	P01103 gallus gall
15	303	19.1	641	1 MYB_CHICK	P46200 bos taurus
16	300	18.9	640	1 MYB_BOVIN	P10244 homo sapien
17	299.5	18.9	700	1 MYBB_HUMAN	P10243 homo sapien
18	296	18.7	451	1 MYBH_DICDI	P34127 dictyostell
19	296	18.7	640	1 MYB_HUMAN	P10242 homo sapien
20	294	18.6	636	1 MYB_MOUSE	P06876 mus musculu
21	291	18.4	751	1 MYBA_MOUSE	P10243 homo sapien
22	288.5	18.2	624	1 MYB_XENLA	P08759 xenopus lae
23	288.5	18.2	752	1 MYBA_HUMAN	P10243 homo sapien
24	284	17.9	757	1 MYBA_CHICK	P52550 gallus gall
25	274	17.3	657	1 MYB_DROME	P04197 drosophil
26	193	12.2	449	1 MYB1_PHYPA	P80074 physcomitre
27	160	10.1	811	1 BAS1_YEAST	P02205 saccharomyc
28	133	8.4	590	1 YM63_YEAST	P03654 saccharomyc
29	132	8.3	757	1 CCS_SCHPO	P39964 schizosacch
30	117	7.4	810	1 REB1_YEAST	P21558 saccharomyc
31	110	6.9	595	1 REB1_KLULA	P03950 kluyveromyc
32	106	6.7	534	1 GCR2_YEAST	P01722 saccharomyc
33	102.5	6.5	661	1 WH13_YEAST	P34761 saccharomyc

34	102.5	6.5	730	1 GLN3_YEAST	P18494 saccharomyc
35	102	6.4	696	1 SKT5_YEAST	P34226 saccharomyc
36	99.5	6.3	1113	1 PER3_MOUSE	P070361 mus musculu
37	98.5	6.2	410	1 LMP2_CRIGR	P49130 cricetus
38	98	6.2	923	1 YEL2_YEAST	P40017 saccharomyc
39	96.5	6.1	652	1 DYN1_DICDI	P54703 dictyostell
40	95.5	6.0	1373	1 HRP1_DICDI	P20425 schizosacch
41	94.5	6.0	781	1 Y3B_YEAST	P3316 saccharomyc
42	94.5	6.0	1704	1 VIR1_FUNHE	P090508 fundulus he
43	93.5	5.9	1142	1 POL_HVZBE	P18093 human immun
44	93.5	5.9	1822	1 YM68_CAEEL	P34529 caenorhabdi
45	92.5	5.8	382	1 MSN1_YEAST	P22148 saccharomyc

## ALIGNMENTS

RESULT	ID	MYB3_MAIZE	STANDARD	PRT	255 AA.
AC	P20025;				
DT	01-FEB-1991 (Rel. 17, Created)				
DT	01-FEB-1991 (Rel. 17, Last sequence update)				
DT	01-OCT-1996 (Rel. 34, Last annotation update)				
DE	MYB-RELATED PROTEIN ZM38.				
OS	Zea mays (Maize).				
OC	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;				
OC	Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC clade;				
OC	Panicoidae; Andropogoneae; Zea.				
OX	NCBI_TaxID=4577;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RX	MEDLINE=89313655; PubMed=2664447;				
RA	Marocco A., Wassenbach M., Becker D., Paz-Ares J., Siedler H.,				
RA	Saladini F., Rohde W.;				
RT	"Multiple genes are transcribed in Hordeum vulgare and Zea mays that				
RT	carry the DNA binding domain of the myb oncoproteins."				
RL	Mol. Gen. Genet. 216:183-187(1989).				
CC	-1- SIMILARITY: BELONGS TO THE MYB FAMILY.				
DR	PIR: S04899; S04899.				
DR	HSSP: P01103; IPOM.				
DR	MalzEDB: 69594; -				
DR	InterPro: IPR001005; MYB_DNA_bind.				
DR	Pfam: PF00249; myb_DNA-binding; 2.				
DR	SMART: SM00395; SANT: 2.				
DR	PROSITE: PS00037; MYB_1; 1.				
DR	PROSITE: PS00334; MYB_2; 1.				
DR	PROSITE: PS50090; MYB_3; 2.				
KW	Nuclear protein; DNA-binding; Repeat; Transcription regulation.				
FT	DNA_BIND 9 61 MYB.				
FT	DNA_BIND 62 102 MYB.				
SO	SEQUENCE 255 AA; 27568 MW; 26AAADCB0633E06 CRC64;				

Query Match	43.8%; Score 693; DB 1; Length 255;
Best Local Similarity	47.1%; Pred. No. 3.4e-46;
Matches 140; Conservative 33; Mismatches 50; Indels 74; Gaps 7;	
1 MGRSPCCERAHNTKGMATKEEDORLIVYRHVSGCWRSLKRAAGLFCGSCRLRWNY	60
	60
1 MGRSPCCERAHNTKGMATKEEDORLIVYRHVSGCWRSLKRAAGLFCGSCRLRWNY	60
	60
61 LRPDLKRGNTAEEDLIVLHSLGKWSLILARLPGRTNETKRYNTHKRLISRG	120
	120
61 LRPDLKRGNTAEEDLIVLHSLGKWSLILARLPGRTNETKRYNTHKRLISRG	120
	120
121 IDPOTHRPNTANTNTVTAPELDFRNPSYVSKSSIKNPISLDFNNEQFKSNTDSL	180
	180
121 IDPOTHRPNTANTNTVTAPELDFRNPSYVSKSSIKNPISLDFNNEQFKSNTDSL	180
	180
181 EEPNCTASSGTTDEEQOELHKKOQYGPSNGDINLELSIGIYSADSRVSNANSAESK	240
	240
181 EEPNCTASSGTTDEEQOELHKKOQYGPSNGDINLELSIGIYSADSRVSNANSAESK	240
	240
161 KAPRC-----PDLNLDLCI-----SPCQOQOEEBEVDK	189



RESULT  
2

ID MYB1\_HORVU STANDARD; PRT; 267 AA.

AC P20026;  
DT 01-FEB-1991 (Rel. 17, Created)  
DT 01-FEB-1991 (Rel. 17, Last sequence update)  
DT 01-OCT-1996 (Rel. 34, Last annotation update)

DE MYB-RELATED PROTEIN HVL.

CN MYB1.

OS Hordeum vulgare (Barley).

OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooidae;  
OC Triticeae; Hordeum.  
OX NCBI\_TaxID=4513;

RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-CV. ABYSSINIAN 2231;  
RX MEDLINE=89313655; PubMed=2664447;  
RA Morocco A., Wiesenbach M., Becker D., Paz-Ares J., Seadler H.,  
RA Salamini F., Rohde W.;  
RT "Multiple genes are transcribed in Hordeum vulgare and Zea mays that  
carry the DNA binding domain of the myb oncoproteins.";  
RL Mol. Gen. Genet. 216:183-187(1989).  
RN [2]

SEQUENCE FROM N.A.  
RC STRAIN-CV. ABYSSINIAN 2231; TISSUE=Leaf;  
RX MEDLINE=94035190; PubMed=8220488;  
RA Wiesenbach M., Ueberlacker B., Vogt F., Becker D., Salamini F.,  
RA Rohde W.;  
RT "Myb genes from Hordeum vulgare: tissue-specific expression of  
chimeric Myb promoter/gus genes in transgenic tobacco.";  
RL Plant J. 4:411-422(1993).

-1 FUNCTION: POSSIBLE TRANSCRIPTION ACTIVATOR IN RESPONSE TO AN  
EXTERNAL SIGNAL. MAY BE INVOLVED IN THE REGULATION OF FLAVONOID  
BIOSYNTHESIS.  
CC -1 SUBCELLULAR LOCATION: NUCLEAR (PROBABLE).  
CC -1 TISSUE SPECIFICITY: GERMINATING SEED AND APICAL MERISTEM OF SHOOT  
AND ROOT.  
CC -1 SIMILARITY: BELONGS TO THE MYB FAMILY.

-----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/  
or send an email to license@isb-sib.ch).

DR EMBL; X70877; CA50222.1; -  
DR EMBL; X70879; CA50224.1; -  
DR PIR; S04896; S04896.  
DR PIR; S31817; S31817.  
DR HSP; P01103; IPOM.  
DR Mendel; 8594; HORVU.MYB.mn8594.  
DR InterPro; IPR001005; MYB\_DNA\_bind.  
DR Pfam; PF00249; myb\_DNA-binding; 2.  
DR SMART; SM00395; SANT; 2.  
DR PROSITE; PS00037; MYB\_1; 1.  
DR PROSITE; PS00334; MYB\_2; 1.  
DR PROSITE; PS50090; MYB\_3; 2.  
KW Nuclear protein; DNA-binding; Activator; Transcription regulation;  
Repeat.  
FT DNAS\_BIND 9 61 MYB.  
FT DNABIND 62 102 MYB.  
SQ SEQUENCE 267 AA; 129740 MW; DECEDCBEA847749 CRC64;

Query Match Summary: 43.1%; Score 683; DB 1; Length 267;  
Best Local Similarity 50.7%; Pred. No. 2, 1e-45;  
Matches 142; Conservative 28; Mismatches 72; Indels 38; Gaps 5;

1 MGSPCCERKANTKAGATKEDORLINTIRYHGECCWRSLPRAAGLLRCGSCRLAMINY 60  
|||||  
1 MGSPCCERKANTKAGATKEDORLINTIRYHGECCWRSLPRAAGLLRCGSCRLAMINY 60  
Db 1 MGSPCCERKANTKAGATKEDORLINTIRYHGECCWRSLPRAAGLLRCGSCRLAMINY 60  
61 LRPDLKRGNTFEDEDELLIKHSLGKNWSLAGLPGRTDDEIKYWNTHIRKLLSRG 120  
|||||  
61 LRPDLKRGNTFEDEDELLIKHSLGKNWSLAGLPGRTDDEIKYWNTHIRKLLSRG 120  
Db 61 LRPDLKRGNTFEDEDELLIKHSLGKNWSLAGLPGRTDDEIKYWNTHIRKLLSRG 120  
121 IDPQTHRLPQNTANTNVTAPTELDERNSPUSKSSSIKNSLSDPNYNEFOFKSNTDSL 180  
|||||  
121 IDPQTHRLPQNTANTNVTAPTELDERNSPUSKSSSIKNSLSDPNYNEFOFKSNTDSL 180  
Db 121 IDPQTHRLPQNTANTNVTAPTELDERNSPUSKSSSIKNSLSDPNYNEFOFKSNTDSL 180  
181 EEFNCTAASSMTTDEEOEQOLHKKOQY---PSNGDINLEISIGIVSADSRSYNSANSA 237  
|||||  
160 PAKAASAAAAASHHVDHHRNSNPOLDWGQGRPKCPDNLDCI-----SPPIHEDPMV 213  
|||||  
238 ESKPKVDNNNFOLFEOAMAKAVCLCW--QLGFGTSEIR 275  
|||||  
214 DFKPVYKR-----EAGVGVGVGGLCFECSMGLPRESSDCK 247  
|||||

RESULT 3  
MB2\_PHYPA  
ID MYB2\_PHYPA STANDARD; PRT; 421 AA.  
AC P80073;  
DT 01-FEB-1994 (Rel. 28, Created)  
DT 01-FEB-1994 (Rel. 28, Last sequence update)  
DT 01-OCT-1996 (Rel. 34, Last annotation update)  
DE MYB-RELATED PROTEIN PP2.  
GN PP2.  
OS Physcomitrella patens (Moss).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Bryophyta;  
OC Bryopsida; Funariidae; Funariales; Funariaceae; Physcomitrella.  
NCBI\_TaxID=3218;  
RN (1)  
RP SEQUENCE FROM N.A.  
RX MEDLINE=94004988; PubMed=8401607;  
RA Leech M.J., Kammerer W., Cove D.J., Martin C., Wang T.L.;  
RT "Expression of myb-related genes in the moss, Physcomitrella patens";  
RL Plant J. 3:51-61(1993).  
CC -1- FUNCTION: POSSIBLE TRANSCRIPTION ACTIVATOR.  
CC -1- SUBCELLULAR LOCATION: NUCLEAR (POTENTIAL).  
CC -1- DEVELOPMENTAL STAGE: HIGH RATES OF GROWTH.  
CC -1- SIMILARITY: BELONGS TO THE MYB FAMILY.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC EMBL: X67050; CAA47435.1; -  
CC DR PIR: S24244; S24244.  
CC DR HSSP: P01103; IPOM.  
CC DR InterPro: IPR001005; Myb-DNA\_Bind.  
CC DR Pfam: PF00249; myb-DNA-binding; 2.  
CC DR SMART, SM00395; SANT; 2.  
CC DR PROSITE: PS00037; MYB\_1; 1.  
CC DR PROSITE: PS00334; MYB\_2; 1.  
CC DR PROSITE: PS50050; MYB\_3; 2.  
CC KW Nuclear protein; DNA-binding; Repeat; Transcription regulation.  
CC FT DNA\_BIND 9 61 MYB.  
CC FT DNA\_BIND 62 102 MYB.  
CC SEQUENCE 421 AA; 46695 MW; 27A357857BD0F3F4 CRC64;

Query Match 32.1%; Score 509; DB 1; Length 421;



Best Local Similarity 55.8%; Pred. No. 7.5e-32;  
Matches 91; Conservative 27; Mismatches 45; Indels 0; Gaps 0;

QY 1 MGRSPCCERAHNKGAWTEEDORLNIYRVHGGCWRSLPKAGLLRCGSCRLRWNY 60  
DB 1 MGRKPCCEKVGKRGWTEEDOKLVSHITNNGSCWRAIPKLGKCGSCRLRWNY 60  
QY 61 LRPLDKRGNTFEEDLITKLSLGNKWSLNGRLPGRTDNEIKNYNTHIRKRLISRG 120  
DB 61 LRPLDKRGNTFEEDLITKLSLGNKWSLNGRLPGRTDNEIKNYNTHIRKRLISRG 120  
QY 121 IDPQTHRLPLQNTANTNTVTAPELDFRNSPTSVSKSSSIXKPS 163  
DB 121 IDPQTHRLPLQNTANTNTVTAPELDFRNSPTSVSKSSSIXKPS 163

RESULT 4  
MYB\_MAIZE  
ID MYB\_MAIZE STANDARD; PRT; 399 AA.  
AC P27898; P27899;  
DT 01-AUG-1992 (rel. 23, Created)  
DT 01-AUG-1992 (rel. 23, Last sequence update)  
DT 15-JUL-1999 (rel. 38, Last annotation update)  
DE MYB-RELATED PROTEIN P.  
GN P.  
OS Zea mays (Maize).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC clade;  
OC Panicoideae; Andropogoneae; Zea.  
OX NCBI\_TaxID=4577;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=91271238; PubMed=2052542;  
RA Groves D. E., Atkinson T. J.;  
RT "Alternatively spliced products of the maize P gene encode proteins with homology to the DNA-binding domain of myb-like transcription factors";  
RT Proc. Natl. Acad. Sci. U.S.A. 88:4587-4591(1991).  
RN [2]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=92275319; PubMed=1317315;  
RA Atkinson T. J., Groves D. E., Peterson T. J.;  
RT "Insertional mutagenesis of the maize P gene by intragenic transposition of Ac";  
RT Genetics 131:199-209(1992).  
CC -1- FUNCTION: TRANSCRIPTION FACTOR POSTULATED TO REGULATE THE BIOSYNTHETIC PATHWAY OF A FLAVONOID-DERIVED PIGMENT IN CERTAIN FLORAL TISSUES.  
CC -1- ALTERNATIVE PRODUCTS: 2 ISOFORMS: A LONG FORM (SHOWN HERE) AND A SHORT FORM; ARE PRODUCED BY ALTERNATIVE SPLICING.  
CC -1- SIMILARITY: BELONGS TO THE MYB FAMILY.  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
DR EMBL: M73028; AAA33500.1; -;  
DR EMBL: M73029; AAA33501.1; -;  
DR EMBL: Z11879; CAA77939.1; -;  
DR PIR: A39697; A39697.  
DR PIR: B39697; B39697.  
DR PIR: S26150; S26150.  
DR HSSP: P01103; 1POM.  
DR TRANSFAC: T01580; -;  
DR TRANSFAC: T01591; -;  
DR MaizeDB: 69180; -;  
DR MaizeDB: 69181; -;  
DR InterPro: IPR001005; MYB\_DNA\_bind.  
DR Pfam: PF00249; myb\_DNA-binding; 2.

DR SMART: SM00395; SANT; 2.  
DR PROSITE: PS00037; MYB\_1; 1.  
DR PROSITE: PS00334; MYB\_2; 1.  
DR PROSITE: PS50090; MYB\_3; 2.  
KW Nuclear protein; DNA-binding; Repeat; Transcription regulation;  
KW Alternative splicing.  
FT DNA\_BIND 9 61  
FT MYB.  
FT DNA\_BIND 62 102  
FT MYB.  
FT DOMAIN 178 181  
FT POLY-ALA.  
FT DOMAIN 305 311  
FT POLY-LEU.  
FT VARSPPLIC 89 152  
FT WSLIASHLPGRTDNEIKNYNTHIRKRLISRG  
FT TAIATDMSKLSQADRRGRTRG -> RHMIRADSPST  
FT VACPLGALATLPLRQSPPTARTITYDRIGSALNLSVRC  
FT FRCVPSRW (IN SHORT ISOFORM).  
FT MISSING (IN SHORT ISOFORM).  
SQ SEQUENCE 399 AA; 43756 MW; E8025B00A4CF5D0 CRC64;

Query Match 31.3%; Score 496.5; DB 1; Length 399;  
Best Local Similarity 42.7%; Pred. No. 6.3e-31;  
Matches 103; Conservative 25; Mismatches 56; Indels 57; Gaps 4;

QY 1 MGRSPCCERAHNKGAWTEEDORLNIYRVHGGCWRSLPKAGLLRCGSCRLRWNY 60  
DB 1 MGRKPCCEKVGKRGWTEEDOKLVSHITNNGSCWRAIPKLGKCGSCRLRWNY 60  
QY 61 LRPLDKRGNTFEEDLITKLSLGNKWSLNGRLPGRTDNEIKNYNTHIRKRLISRG 118  
DB 61 LRPLDKRGNTFEEDLITKLSLGNKWSLNGRLPGRTDNEIKNYNTHIRKRLISRG 118  
QY 121 IDPQTHRLPLQNTANTNTVTAPELDFRNSPTSVSKSSSIXKPS 153  
DB 121 IDPQTHRLPLQNTANTNTVTAPELDFRNSPTSVSKSSSIXKPS 153  
QY 154 SKS---SSIKNSLDFNNEFQKSTSLDEPNCASSGMTDEEOQLHKKQYGPS 210  
DB 154 SKS---SSIKNSLDFNNEFQKSTSLDEPNCASSGMTDEEOQLHKKQYGPS 210  
QY 181 ASSPRSDVNVNPGPN-----QNSSSGSGTGAEE-----GPS 213  
DB 181 ASSPRSDVNVNPGPN-----QNSSSGSGTGAEE-----GPS 213  
QY 211 N 211  
DB 211 S 214

RESULT 5  
MYB\_MAIZE  
ID MYB\_MAIZE STANDARD; PRT; 340 AA.  
AC P20024;  
DT 01-FEB-1991 (rel. 17, Created)  
DT 01-FEB-1991 (rel. 17, Last sequence update)  
DT 01-OCT-1996 (rel. 34, Last annotation update)  
DE MYB-RELATED PROTEIN ZML.  
OS Zea mays (Maize).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC clade;  
OC Panicoideae; Andropogoneae; Zea.  
OX NCBI\_TaxID=4577;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=89313655; PubMed=2664447;  
RA Marocco A., Wiesenbach M., Becker D., Paz-Ares J., Saedler H.,  
RT "Multiple genes are transcribed in Hordeum vulgare and Zea mays that carry the DNA binding domain of the myb oncoproteins";  
RT Mol. Gen. Genet. 216:183-187(1989).  
CC -1- SIMILARITY: BELONGS TO THE MYB FAMILY.  
CC PIR: S04898; S04898.  
DR HSSP: P06876; 1MBJ.  
DR MaizeDB: 69590; -;  
DR InterPro: IPR001005; MYB\_DNA\_bind.  
DR Pfam: PF00249; myb\_DNA-binding; 2.  
DR SMART: SM00395; SANT; 2.  
DR PROSITE: PS00037; MYB\_1; 1.  
DR PROSITE: PS00334; MYB\_2; 1.



```

DR MalzEDB:24964; -
DR InterPro: IPR001005; Myb_DNA_bind.
DR Pfam: PF00249; myb_DNA-binding; 2.
DR SMART: SM00395; SANT; 2.
DR PROSITE: PS00037; MYB_1; 1.
DR PROSITE: PS00334; MYB_2; 1.
DR PROSITE: PS50090; MYB_3; 2.
KW Nuclear protein; DNA-binding; Repeat; Transcription regulation;
KW Activator; Trans-acting factor.
FT DNA_BIND 9 61 MYB.
FT DNA_BIND 62 102 MYB.
FT DOMAIN 203 212 POLY-GLY.
SO SEQUENCE 252 AA; 26419 MW; AD61FA2B829C5FBD CRC64;

Query Match 27.7%; Score 439; DB 1; Length 252;
Best Local Similarity 63.5%; Pred. No. 8, 8e-27;
Matches 73; Conservative 20; Mismatches 22; Indels 0; Gaps 0;

OY 1 MGRSPCCERAKHTNKAMTKEEDRLINIRYHGEGCWSRLPRAAGILRGKSGCRLMWY 60
DB 1 MGRRACCAKCGVRCAGMTSKEDDALAAVYKHAHGEGKRWREVPKAGILRGKSGCRLMWY 60
OY 61 LRPDLKRGNTFEDEDEILIKHLSLGNKWSLACGLPGCTDNDIKYNTHTIRK 115
DB 61 LRPNIRGNISYDEEDLILRLRLGLNWSLACGLPGCTDNDIKYNTHTIRK 115

RESULT 9
G1L ARATH STANDARD; PRT; 228 AA.
ID G1L ARATH AC P27900;
DT 01-AUG-1992 (Rel. 23, Created)
DT 01-AUG-1992 (Rel. 23, Last sequence update)
DT 01-OCT-1996 (Rel. 34, Last annotation update)
DE TRICHOME DIFFERENTIATION PROTEIN G1L.
GN G1L.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsids.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=92034971; PubMed=1934056;
RA Oppenheimer D.G., Herman P.L., Sivakumaran S., Esch J., Marks M.D.;
RT "A myb gene required for leaf trichome differentiation in Arabidopsis
RT is expressed in stipules."
RL Cell 67:483-492(1991).
CC -1- FUNCTION: REGULATES THE PRODUCTION OF A SIGNAL THAT INDUCES
CC HAIR (TRICHOME) PRECURSOR CELLS ON LEAF PRIMORDIA TO
CC DIFFERENTIATE
CC -1- SIMILARITY: BELONGS TO THE MYB FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
DR EMBL: M79448; AAC97387.1; -
DR PIR: A39289; TYMU01.
DR HSSP: P01103; IPOM.
DR TRANSFAC: T01588; -
DR InterPro: IPR001005; Myb_DNA_bind.
DR Pfam: PF00249; myb_DNA-binding; 2.
DR SMART: SM00395; SANT; 2.
DR PROSITE: PS00037; MYB_1; 1.
DR PROSITE: PS00334; MYB_2; 1.
DR PROSITE: PS50090; MYB_3; 2.
KW Nuclear protein; DNA-binding; Repeat; Transcription regulation.

```

FT DNA\_BIND 11 63 MYB.  
 FT DNA\_BIND 64 104 MYB.  
 SQ SEQUENCE 228 AA: 26435 MW: 8236EA36FAA6B6C CRC64;

Query Match 25.4%; Score 402.5; DB 1; Length 228;  
 Best Local Similarity 37.7%; Pred. No. 4,8e-24;  
 Matches 89; Conservative 32; Mismatches 46; Indels 69; Gaps 7;

DB 8 EKATNKAMTKEDORLINTRVHGECCWRSIPKAAGLRCCKSLRNINIRPDLKR 67  
 10 ENOEYKKGIMVVEENIIMADYLNHTGQMNIRVKTGLKRCCKSCLRMNNTYSPVNVK 69  
 68 GNFTDEEDLIRKLKSLGNKMSLGRPLGRTDNEIKNWNTHIRKLLSRGIDPQTHR 127  
 70 GNFTDEEDLIRKLKSLGNKMSLGRPLGRTDNOYKNWNTHLSKLVG----- 120  
 DB 128 PLNQNTANTNTVAPTELPFRNSPTSVSKSSSIKNPSLDFNYNEFOKSNMDSLEPNCCTA 187  
 121 -----DY-----SSAVKTTGED-----DDSPSLFITA 143  
 QY 188 SSGMTTDEEOQLHKQOQPSNGODINLELSI-GIVSADSRVSNANSAESKPK 242  
 144 AT-PSSRHQDENIYE-----NIAKSFNGVSA-----SYEDKPK 177

RESULT 10  
 MYB\_AVIMB STANDARD; PRT; 382 AA.

AC P01104; Q83871; Q90954;  
 DT 21-JUL-1986 (Rel. 01, Created)  
 DT 01-MAR-1989 (Rel. 10, Last sequence update)  
 DT 20-AUG-2001 (Rel. 40, Last annotation update)  
 DE MYB TRANSFORMING PROTEIN.  
 GN V-MYB.  
 OS Avian myeloblastosis virus.  
 OC Viruses; Retroviridae; Avian type C retroviruses.  
 NCBI\_TaxID=11866;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA MEDLINE=83129359; PubMed=6297766;  
 RA Klempauer K.-H., Gonda T.J., Bishop J.M.;  
 RT "Nucleotide sequence of the retroviral leukemia gene v-myb and its cellular progenitor c-myb: the architecture of a transduced oncogene";  
 RT Cell 31:453-463(1982).  
 RL [2]  
 RN SEQUENCE FROM N.A.  
 RP MEDLINE=8223743; PubMed=6283631;  
 RA Rushlow K.E., Lautenberger J.A., Papas T.S., Baluda M.A., Ferbal B.,  
 RA "Nucleotide sequence of the transforming gene of avian myeloblastosis virus";  
 RT Science 216:1421-1423(1982).  
 RL [3]  
 RN SEQUENCE FROM N.A.  
 RP MEDLINE=94366722; PubMed=8084581;  
 RA Baluda M.A., Reddy E.P.;  
 RT "Anatomy of an integrated avian myeloblastosis provirus: structure and function";  
 RT Oncogene 9:2761-2774(1994).  
 RL [4]  
 RN THE SEQUENCE YAAC(G/T)G  
 CC -1- DOMAIN: THE ABSENCE OF THE MYB N- AND/OR C-TERMINAL OF THIS  
 CC PROTEIN MAY PREVENT HOMODIMER FORMATION, LEADING TO DNA-BINDING  
 CC AND ITS INCREASED TRANSACTIVATION ACTIVITY.  
 CC -1- DISEASE: THE V-MYB ONCOGENE TRANSFORMS IMMATURE MYELOMONOCYTIC  
 CC AVIAN CELLS IN CULTURE AND INDUCES MYELOBLASTOSIS (MYELOID  
 CC LEUKEMIA) IN CHICKENS.  
 CC -1- SIMILARITY: BELONGS TO THE MYB FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC-  
 DR EMBL: J02012; AAA42553.1; ALT\_INIT.  
 DR EMBL: S74099; AAB31930.2;  
 DR EMBL: J02013; AAC15907.1; ALT\_INIT.  
 DR PIR: A01347; QOYV.  
 DR HSP: P06876; IMB.  
 DR TRANSFAC: T00895;  
 DR InterPro: IPR001005; Myb\_DNA\_bind.  
 DR Pfam: PF00249; myb\_DNA-binding; 3.  
 DR SMART: SM00395; SANT; 2.  
 DR PROSITE: PS00037; MYB\_1; 2.  
 DR PROSITE: PS00334; MYB\_2; 2.  
 DR PROSITE: PS50090; MYB\_3; 2.  
 DR TRANSFORMING PROTEIN; Oncogene; Nuclear protein; DNA-binding; Repeat.  
 FT DNA\_BIND 1 15  
 FT DNA\_BIND 16 67 MYB.  
 FT DNA\_BIND 68 118 MYB.  
 FT DNA\_BIND 204 236 MYB (PARTIAL).  
 FT DOMAIN 257 376  
 FT DOMAIN 305 326  
 FT DOMAIN 382 43061 MW: 1459BACAD29EC971 CRC64;  
 FT LEUCINE-ZIPPER.  
 FT NEGATIVE REGULATORY DOMAIN (BY  
 FT SIMILARITY) (INCOMPLETE).  
 FT SEQUENCE

Query Match 20.3%; Score 322; DB 1; Length 382;  
 Best Local Similarity 28.9%; Pred. No. 1.3e-17;  
 Matches 84; Conservative 49; Mismatches 112; Indels 46; Gaps 7;

QY 13 NKGAMTKEDORLINTRVHGECCWRSIPK-AAGLLRCCKSLRNINIRPDLKAGNT 71  
 20 NKQPMTEKEQDQRYVEHQVOKGPKRMSDIAKHLKG--RIGKOCERHWNHNLNPEVKTSTW 77  
 DB 72 EEDDELIRKLSLGNKMSLGRPLGRTDNEIKNWNTHIRKLLSRGIDPQTHRPLNO 131  
 78 EEDDRITLYQAHKRLGNMAELAKLLPGRTDNAVKNHNSMTRKRYEQEYPOESSKAGRP 137  
 QY 132 TANTNTVAPTELPFRNSPTSVSKSSSIKNP-SLDFNYNEFOKSNMDSLEPNCCTA 178  
 DB 138 SATTGQKSSHLMFAFNPNRPGPLGAGAPLGSDPYHYHIAEPQVPCQIPYVALHIN 197  
 QY 179 --SLEPNCYASSGMTTDEEOQDLH-----KKOQYSPNGODINLELSI 221  
 DB 198 IINVPQPAALAIORHYDEDEPEKERIKELLLMSTENELKGOALPPOHNTANYPGW 257  
 QY 222 GIVSADSRVSNANSAES-----KPKVDNNFQLEQAMVAKAVCL 262  
 DB 238 STTVADNTRISGDNAVSCLEHHHCITSPYVDHG---CLPESSASPAKCM 305

RESULT 11  
 MYBB\_CHICK STANDARD; PRT; 686 AA.

AC Q03257;  
 DT 01-OCT-1993 (Rel. 27, Created)  
 DT 01-OCT-1993 (Rel. 27, Last sequence update)  
 DT 15-JUL-1998 (Rel. 36, Last annotation update)  
 DE MYB-RELATED PROTEIN B (B-MYB).  
 GN MYB2 OR BMYB.  
 OS Gallus gallus (Chicken).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;  
 OC Gallus.  
 NCBI\_TaxID=9031;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA MEDLINE=93049214; PubMed=1425593;  
 RA Foos G., Grimm S., Klempauer K.-H.;

\*functional antagonism between members of the myb family: B-myb inhibits v-myb-induced gene activation.\*  
 RT EMBL J. 11:4619-4629(1992).  
 RN [2]  
 RP STRUCTURE BY NMR OF 79-186.  
 RX MEDLINE-98322063; PubMed-9657674;  
 RA McInosh P.B., Frenkel T.A., Wollborn U., McCormick J.E.,  
 RA Klemmner K.H., Feeney J., Carr M.D.;  
 RT "Solution structure of the B-myb DNA-binding domain: a possible role for conformational instability of the protein in DNA binding and control of gene expression.";  
 RL Biochemistry 37:9619-9629(1998).  
 CC -1- FUNCTION: IT REPRESENTS V-MYB- AND C-MYB-MEDIATED ACTIVATION OF THE MIM-1 GENE, PROBABLY BY COMPETING WITH OTHER MYB PROTEINS FOR BINDING SITES. IT IS AN INHIBITORY MEMBER OF THE MYB FAMILY.  
 CC -1- SUBCELLULAR LOCATION: NUCLEAR.  
 CC -1- TISSUE SPECIFICITY: EXPRESSED IN HEMATOPOIETIC AND NON HEMATOPOIETIC CELLS.  
 CC -1- SIMILARITY: BELONGS TO THE MYB FAMILY.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC EMBL: X67505; CAA47839.1; -  
 DR PIR: S28050; S28050.  
 DR PDB: 1A57; 01-JUL-98.  
 DR TRANSEAC: T01586; -  
 DR InterPro: IPR001005; Myb\_DNA\_bind.  
 DR Pfam: PF00249; myb-DNA-binding; 3.  
 DR SMART: SM00395; SANT; 3.  
 DR PROSITE: PS00037; MYB\_1; 3.  
 DR PROSITE: PS00034; MYB\_2; 3.  
 DR PROSITE: PS50090; MYB\_3; 3.  
 KW Transcription regulation; Nuclear protein; DNA-binding; Repeat; 3D-structure.  
 KM DNA\_BIND 25 77 MYB.  
 FT DNA\_BIND 78 129 MYB.  
 FT DNA\_BIND 130 180 MYB.  
 FT SEQUENCE 686 AA; 77736 MW; 2D1209EAD6489D7B CRC64;  
 SO  
 Query Match 19.7%; Score 312; DB 1; Length 686;  
 Best Local Similarity 38.6%; Pred. No. 1.6e-16;  
 Matches 64; Conservative 33; Mismatches 59; Indels 10; Gaps 3;  
 QY 14 KGAWTKEDDRLNIRVHGGCWRSLPK--AAGLLRCGSKRLRWNYLRPDLKRGNTFTE 72  
 DB 83 KGPWTKEDDQVIELVHKYGRKQWTLAKHLKG--RLGKQCRERHNLNLPDYKSSWTE 140  
 QY 73 EEDELIIKLSLGNKWSLILAGRLPGRTDNEIKRWYTHIRKILSRIDPQTHRLPLNOT 132  
 DB 141 EEDRIIISAKHMGNSLAKLILGRDIDNAKHNHNSITIKRKYDVGCF-----LNET 193  
 QY 133 ANNTNTVTAPELDFRNSPTSVSKSSIKNPSLDFNYNEFOKSTDSLEBN 178  
 DB 194 KESQPLVLLVEVDNDSQSGTRASQTIIVNMPVDISEIKEDVDSD 239  
 RESULT 12  
 ID MYB\_XENLA STANDARD; PRT: 728 AA.  
 AC 005935;  
 DT 01-OCT-1996 (Rel. 34, Created).  
 DT 01-OCT-1996 (Rel. 34, Last sequence update).  
 DT 01-OCT-1996 (Rel. 34, Last annotation update).  
 DE MYB-RELATED PROTEIN A (A-MYB) (XANTYB) (MYB-RELATED PROTEIN 2) (XMYB2).  
 GN MYB1 OR AMYB OR MYB2.  
 OS Xenopus laevis (African clawed frog);

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 CC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae;  
 CC Xenopodinae; Xenopus.  
 OX NCBI\_Taxid-8355;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE-93284408; PubMed-8510936;  
 RA Sleeman J.P.;  
 RT "Xenopus A-myb is expressed during early spermatogenesis.";  
 RL Oncogene 8:1931-1941(1993).  
 RN [2]  
 RP SEQUENCE OF 1-181 FROM N.A.  
 RX MEDLINE-92297434; PubMed-1606020;  
 RA Bouwmeester T., Guemann S., El-Baradi T., Kalkbrenner F.,  
 RA van Wijk I., Moelling K., Pieler T.;  
 RT "Molecular cloning, expression and in vitro functional characterization of Myb-related proteins in Xenopus.";  
 RL Mech. Dev. 37:57-68(1992).  
 CC -1- FUNCTION: STRONG TRANSCRIPTIONAL ACTIVATOR; DNA-BINDING PROTEIN THAT SPECIFICALLY RECOGNIZES THE SEQUENCE 5'-YAC(G/T)G-3'. COULD HAVE A ROLE IN THE PROLIFERATION AND/OR DIFFERENTIATION OF NEUROGENIC, SPERMATOGENIC AND B-LYMPHOID CELLS (BY SIMILARITY).  
 CC -1- SUBCELLULAR LOCATION: NUCLEAR.  
 CC -1- DEVELOPMENTAL STAGE: EXPRESSED DURING EARLY SPERMATOGENESIS.  
 CC -1- SIMILARITY: BELONGS TO THE MYB FAMILY.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC EMBL: X72620; CAA51196.1; -  
 DR EMBL: M75871; AAA49904.1; -  
 DR HSRP: P06876; 1MBJ.  
 DR InterPro: IPR001005; Myb\_DNA\_bind.  
 DR Pfam: PF00249; myb-DNA-binding; 3.  
 DR SMART: SM00395; SANT; 3.  
 DR PROSITE: PS00037; MYB\_1; 3.  
 DR PROSITE: PS00034; MYB\_2; 2.  
 DR PROSITE: PS50090; MYB\_3; 3.  
 KW Transcription regulation; Activator; Nuclear protein; DNA-binding; Repeat.  
 KM DNA\_BIND 29 80 MYB.  
 FT DNA\_BIND 81 132 MYB.  
 FT DNA\_BIND 133 183 MYB.  
 FT DOMAIN 230 293  
 FT DOMAIN 296 534  
 FT SEQUENCE 728 AA; 83576 MW; 7A360CBA51AC0797 CRC64;  
 SO  
 Query Match 19.5%; Score 309; DB 1; Length 728;  
 Best Local Similarity 29.0%; Pred. No. 2.9e-16;  
 Matches 83; Conservative 45; Mismatches 96; Indels 62; Gaps 10;  
 QY 14 KGAWTKEDDRLNIRVHGGCWRSLPK--AAGLLRCGSKRLRWNYLRPDLKRGNTFTE 72  
 DB 86 KGPWTKEDDQVIELVHKYGRKQWTLAKHLKG--RLGKQCRERHNLNLPDYKSSWTE 143  
 QY 73 EEDELIIKLSLGNKWSLILAGRLPGRTDNEIKRWYTHIRKILSRIDPQTHRLPLNOT 129  
 DB 144 EEDRIIISAKHMGNSLAKLILGRDIDNAKHNHNSITIKRKYDVGCF-----LNET 193  
 QY 130 ANNTNTVTAPELDFRNSPTSVSKSSIKNPSLDFNYNEFOKSTDSLEBN 184  
 DB 204 KIQA--KSCAAPNHLDAQNFYIPVQTOIRYSSLSHD-----NCTIIONSFSTIQDPF 256  
 QY 185 CASSGNTDEQOELHKKQOYGPSNGODINIELSIGIVASDSSRVANSASEKPKVD 244

DB 257 VDA-----DDEPKERRIK-----ELELLLSAENE----- 281

OY 245 NNNFOLEQAMVAKAVCLCWLQGTSEICRNCN---SNSNGFYS 287

DB 282 -----VRRKRVPSSSLTMSSESYHMGESMNTMSHLEQCHDFYS 321

RESULT 13

ID MYB\_XENLA STANDARD: PRT: 743 AA.

AC P52551:

DT 01-OCT-1996 (Rel. 34, Created)

DT 30-MAY-2000 (Rel. 39, Last sequence update)

DT 30-MAY-2000 (Rel. 39, Last annotation update)

DE MYB-RELATED PROTEIN B (B-MYB) (MYB-RELATED PROTEIN 1) (XMYB1).

GN MYB12 OR BMYB OR MYB1.

OS Xenopus laevis (African clawed frog).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae;

OC Xenopodidae; Xenopus.

OX NCBI\_TaxID=83355;

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE=92297434; PubMed=1606020;

RA Bouwmeester T., Guenham S., El-Baradi T., Kalkbrenner F.,

RA van Wijk I., Moelling K., Pieler T.;

RT "Molecular cloning, expression and in vitro functional

RT characterization of Myb-related proteins in Xenopus.";

RL Mech. Dev. 37:57-68(1992).

RN [2]

RP REVISIONS TO C-TERMINUS.

RA Humbert-Lan G., Pieler T.;

RT "Regulation of DNA-binding activity and nuclear transport of B-Myb in

RT Xenopus oocytes.";

RL Submitted (DEC-1998) to the EMBL/GenBank/DBJ databases.

CC -1- SUBCELLULAR LOCATION: NUCLEAR.

CC -1- DEVELOPMENTAL STAGE: PRESENT THROUGHOUT OOGENESIS AND EARLY

CC XENOPUS EMBRYOGENESIS; IN ADULT TISSUE IT IS PRIMARILY DETECTED IN

CC BLOOD.

CC -1- SIMILARITY: BELONGS TO THE MYB FAMILY.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC -----

CC EMBL: M75870; AAC98701.1; -

DR HSSP: 003237; IASJ.

DR InterPro: IPR001005; Myb\_DNA\_bind.

DR Pfam: PF00249; myb\_DNA\_binding; 3.

DR SMART: SM00395; SANT; 3.

DR PROSITE: PS00037; MYB\_1; 3.

DR PROSITE: PS00334; MYB\_2; 3.

DR PROSITE: PS50090; MYB\_3; 3.

KW Transcription regulation; Nuclear protein; DNA-binding; Repeat.

FT DNA\_BIND 25 77 MYB.

FT DNA\_BIND 78 129 MYB.

FT DNA\_BIND 130 180 MYB.

SO SEQUENCE 743 AA; 82909 MW; 5FD1D678BB24409B CRC64;

Query Match 19 38; Score 306.5; DB 1; Length 743;

Best Local Similarity 30.88; Pred. NO. 4.6e-16;

Matches 80; Conservative 37; Mismatches 56; Indels 87; Gaps 9;

OY 14 KGAWTKEDQRLINIVHSEGGCRSLPK-AAGLLRCGSCRLMIVNLPDLKGNFTE 72

DB 83 KGPWTKEDQRLINIVHSEGGCRSLPK-AAGLLRCGSCRLMIVNLPDLKGNFTE 72

OY 73 EDELLIKHSLGKWSLTAGRLPGRTDNEIKRWMTHTIRKLISRG1----- 121

DB 141 EEDRIIOAHNVLDGNRMALIKLLPGRTDNVAKNMHNSITKRKVEGGLTVAKSGOEE 200

OY 122 -----DPQTH-----RPLNQTANTMTVAPTELDFRNSPTSVKSSIRKNSLDENY 168

DB 201 REDSGYQAEQGNHVLSEPPERSAN-----IPEE-----PSNLSPKL----- 239

OY 169 NEFOKSNSTSLSEPNCTASSGMDTEQEQQLHKKQOYGPSNGODINLELSIGVPSADS 228

DB 240 -----LTKSPQIRSEOE-----SGEGSNSSESAIY----- 266

OY 229 SRVSNANSAESKPKVDNNNF 248

DB 267 -----DSAPKRWVEYVNF 280

RESULT 14

ID MYB\_MOUSE STANDARD: PRT: 704 AA.

AC P48972;

DT 01-FEB-1996 (Rel. 33, Created)

DT 01-FEB-1996 (Rel. 33, Last sequence update)

DT 20-AUG-2001 (Rel. 40, Last annotation update)

DE MYB-RELATED PROTEIN B (B-MYB).

GN MYB12 OR BMYB.

OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Sclurognathi; Muridae; Murinae; Mus.

OX NCBI\_TaxID=10090;

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE=92366176; PubMed=1501895;

RA Lam E.W., Robinson C., Watson R.J.;

RT "Characterization and cell cycle-regulated expression of mouse

RT B-myb.";

RL Oncogene 7:1885-1890(1992).

RN [2]

RP SEQUENCE OF 1-6 FROM N.A.

RX MEDLINE=93327760; PubMed=8334989;

RA Lam E.W., Watson R.J.;

RT "An E2F-binding site mediates cell-cycle regulated repression of

RT mouse B-myb transcription.";

RL EMBO J. 12:2705-2713(1993).

CC -1- SUBCELLULAR LOCATION: NUCLEAR.

CC -1- PTM: PHOSPHORYLATED BY CYCLIN A/CDK2 DURING S-PHASE.

CC PHOSPHORYLATION AT THR-524 IS PROBABLY INVOLVED IN TRANSCRIPTIONAL

CC ACTIVITY (BY SIMILARITY).

CC -1- SIMILARITY: BELONGS TO THE MYB FAMILY.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC -----

CC EMBL: X70472; CAA49898.1; -

DR EMBL: X73028; CAA51511.1; -

DR HSSP: 003237; IASJ.

DR MGD: MGI:101785; MYB12.

DR InterPro: IPR001005; Myb\_DNA\_bind.

DR Pfam: PF00249; myb\_DNA\_binding; 3.

DR SMART: SM00395; SANT; 3.

DR PROSITE: PS00037; MYB\_1; 3.

DR PROSITE: PS00334; MYB\_2; 3.

DR PROSITE: PS50090; MYB\_3; 3.

KW Transcription regulation; Nuclear protein; DNA-binding; Repeat;

KW Phosphorylation.

FT DNA\_BIND 25 77 MYB.

FT DNA\_BIND 78 129 MYB.

FT DNA\_BIND 130 180 MYB.

FT MOD\_RES 443 443 PHOSPHORYLATION (BY CDK2) (BY







GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 23, 2002, 21:11:38 ; Search time 91.99 Seconds

(without alignments)  
467.486 Million cell updates/sec

Title: US-09-453-387A-2  
Perfect score: 1584  
Sequence: 1 MGRSPCCERAHNTKAWTKE.....RNCNSNSNGFYSCRPIDS 294

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 473505 seqs, 146272329 residues  
Total number of hits satisfying chosen parameters: 473505

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :  
1: SP:REMBL\_17:\*  
2: SP:archaea:\*  
3: SP:fungi:\*  
4: SP:human:\*  
5: SP:invertebrate:\*  
6: SP:mammal:\*  
7: SP:mhc:\*  
8: SP:organelle:\*  
9: SP:phage:\*  
10: SP:plant:\*  
11: SP:rodent:\*  
12: SP:virus:\*  
13: SP:vertebrate:\*  
14: SP:unclassified:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1584	100.0	294	10	043436 gossypium h
2	854.5	53.9	274	10	P81395 arabidopsis
3	768	48.5	257	10	Q959K9 arabidopsis
4	762	48.1	236	10	Q38851 arabidopsis
5	756.5	47.8	251	10	Q06944 hordeum vul
6	728	46.0	273	10	Q43525 lycopersico
7	726	45.8	264	10	Q9ATD1 gossypium h
8	723	45.6	282	10	Q9SZP1 arabidopsis
9	717	45.3	282	10	Q49774 arabidopsis
10	704.5	44.5	232	10	P81393 antirrhinum
11	702	44.3	246	10	Q957Y2 arabidopsis
12	684	43.2	274	10	Q9SM28 arabidopsis
13	684	43.2	274	10	Q49608 arabidopsis
14	683	43.1	239	10	Q23889 oryza sativ
15	681.5	43.0	269	10	Q42379 arabidopsis
16	674.5	42.6	212	10	Q9SD58 arabidopsis
17	652	41.2	276	10	Q23890 oryza sativ
18	626.5	39.6	226	10	Q49021 gossypium h
19	584.5	36.9	321	10	Q9LID5 arabidopsis

20	584.5	36.9	352	10	Q9LDE2	Q9LDE2 arabidopsis
21	567	35.8	325	10	Q9LEF1	Q9LEF1 arabidopsis
22	553.5	34.9	345	10	Q9LE63	Q9LE63 arabidopsis
23	549.5	34.7	282	10	Q9M0J5	Q9M0J5 arabidopsis
24	546.5	34.5	274	10	Q40173	Q40173 lycopersico
25	546	34.5	247	10	Q9ATD9	Q9ATD9 gossypium h
26	543	34.3	414	10	Q9SAV9	Q9SAV9 arabidopsis
27	539.5	34.1	254	10	Q9ATD2	Q9ATD2 gossypium h
28	539.5	34.0	350	10	Q9LDR8	Q9LDR8 arabidopsis
29	539	34.0	352	10	Q22685	Q22685 arabidopsis
30	537	33.9	368	10	Q04141	Q04141 oryza sativ
31	536	33.8	388	10	Q40920	Q40920 picea maria
32	534.5	33.7	324	10	Q9M0Y5	Q9M0Y5 arabidopsis
33	533.5	33.7	286	10	Q9LXF1	Q9LXF1 arabidopsis
34	531.5	33.6	326	10	Q9SP61	Q9SP61 arabidopsis
35	531	33.5	327	10	Q9SP67	Q9SP67 arabidopsis
36	530.5	33.5	269	10	Q9SE64	Q9SE64 plumbella
37	530	33.5	282	10	Q9CTU7	Q9CTU7 arabidopsis
38	528.5	33.4	352	10	Q65409	Q65409 arabidopsis
39	528	33.3	421	10	Q02992	Q02992 petunia hyb
40	526	33.2	314	10	Q9C695	Q9C695 arabidopsis
41	526	33.2	578	10	Q9FVU6	Q9FVU6 arabidopsis
42	524.5	33.1	340	10	Q9FR07	Q9FR07 zea mays su
43	524	33.1	249	10	Q38850	Q38850 arabidopsis
44	523.5	33.0	336	10	Q9FLR1	Q9FLR1 arabidopsis
45	522	33.0	299	10	Q9M2D9	Q9M2D9 arabidopsis

## ALIGNMENTS

## RESULT 1

ID Q43436 PRELIMINARY; PRT; 294 AA.

AC Q43436;  
DT 01-NOV-1996 (TREMBLrel. 01, Created)  
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)  
DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)  
DE COPTON DNA-BINDING DOMAIN MRNA.  
GN MYB.  
OS Gossypium hirsutum (Upland cotton).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons; rosidae;  
OC eurosids II; Malvales; Malvaceae; Gossypium.  
OX NCBI\_TaxID=3635;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=CV. ACALA SJ-2; TISSUE=OVULE;  
RA Wilkins T.A., Lu C.C.;  
RL Submitted (JAN-1994) to the EMBL/GenBank/DBJ databases.  
DR EMBL; L04497; AAA33067.1; .  
DR HSSP; P01103; IPOM.  
DR TRANSFAC; T02864; .  
DR Mendel; 8407; Goshi.Myb.8407.  
DR InterPro; IPR001005; MYB\_DNA\_bind.  
DR Pfam; PF00249; MYB\_DNA-binding; 2.  
DR SMART; SM00395; SANT; 2.  
DR PROSITE; PS00037; MYB\_1; UNKNOWN\_1.  
DR PROSITE; PS00334; MYB\_2; 1.  
DR PROSITE; PS50090; MYB\_3; 2.  
KW DNA-binding.  
SQ SEQUENCE 294 AA; 33202 MW; EF1A55B64919F0AC CRC64;

Query Match 100.0%; Score 1584; DB 10; Length 294;

Best Local Similarity 100.0%; Pred. No. 2.2e-130; Gaps 0;

Matches 294; Conservative 0; Mismatches 0; Indels 0;

OY 1 MGRSPCCERAHNTKAWTKEEDORLNTYFVHGEGCRSLPKAGILRCGSCRLRWNY 60  
DB 1 MGRSPCCERAHNTKAWTKEEDORLNTYFVHGEGCRSLPKAGILRCGSCRLRWNY 60  
OY 61 LRPDLKRGNTFEDELLIKLHSLGKWKSLIAGRLPGRTDNEIKNYWTHIKRKLISRG 120

```

Db 61 LRPDLKRGNTFEEDELLIKLHSLGKWSLIAGRLGRDNEIKYWNTHIKRKLISRG 120
OY 121 IDPQTHRLPNOTANTNTVTAPELDFRNSPTSVSKSSIKRPSLDENFNEOFKNTDSL 180
Db 121 IDPQTHRLPNOTANTNTVTAPELDFRNSPTSVSKSSIKRPSLDENFNEOFKNTDSL 180
OY 181 EEPNCTASSGMTTDEEOQLHKKOQYGPNSGODINLELSIGIVSADSSRYSNANSASRK 240
Db 181 EEPNCTASSGMTTDEEOQLHKKOQYGPNSGODINLELSIGIVSADSSRYSNANSASRK 240
OY 241 PRVDNNNFQFLDAQAMAKAVCLCWLQGFSTSEICRNCNSNSNGFYSICRPLDS 294
Db 241 PRVDNNNFQFLDAQAMAKAVCLCWLQGFSTSEICRNCNSNSNGFYSICRPLDS 294

```

## RESULT 2

```

P81395 PRELIMINARY; PRT; 274 AA.
AC P81395;
DT 01-JUN-1998 (TREMBlrel. 06, Created)
DT 01-JUN-1998 (TREMBlrel. 06, Last sequence update)
DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)
DE MYB-RELATED PROTEIN 330.
GN MYB 330.
OS Antirrhinum majus (Garden snapdragon).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC Asteridae; euasterids I; Lamiales; Veroniceae; Antirrhinum.
OX NCBI_TaxID=4151;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-J1:522; TISSUE=FLOWER BUDS;
RX MEDLINE=93005689; PubMed=1840903;
RA Jackson D., Callan-Mackay F., Prescott A.G., Roberts K., Martin C.;
RA "Expression patterns of myb genes from Antirrhinum flowers.";
RL Plant Cell 3:115-125(1991).
CC -1- FUNCTION: MAY BE A TRANSCRIPTIONAL ACTIVATOR.
CC -1- TISSUE SPECIFICITY: MATURE FLOWER.
CC -1- SIMILARITY: BELONGS TO THE MYB FAMILY OF PROTEINS.
DR PIR: J00957; J00957.
DR HSSP: P01103; IPOK.
DR TRANSFAC: T02928; -.
DR Mendel: 24823; Antma.Myb.24823.
DR InterPro: IPR001005; Myb_DNA_bind.
DR Pfam: PF00249; myb_DNA-binding; 2.
DR SMART: SM00395; SANT; 2.
DR PROSITE: PS00037; MYB_1; UNKNOWN_1.
DR PROSITE: PS00334; MYB_2; 1.
DR PROSITE: PS00090; MYB_3; 2.
DR Nuclear protein; DNA-binding; Repeat; Transcription regulation.
FT DNA_BIND 9 61 MYB.
FT DNAS_BIND 62 102 MYB.
SQ SEQUENCE 274 AA; 30679 MW; 69831BC41161BEBB CRC64;

```

Query Match 53.9%; Score 854.5; DB 10; Length 274;  
 Best Local Similarity 60.5%; Pred. No. 8.6e-67;  
 Matches 178; Conservative 23; Mismatches 68; Indels 25; Gaps 8;

```

OY 1 MGRSPCCERKANTNKGAWKEDQRLINIRVHGECWRSLEPKAAGLLRCGSKRLRWNY 60
Db 1 MGRSPCCERKANTNKGAWKEDQRLINIRVHGECWRSLEPKAAGLLRCGSKRLRWNY 60
OY 61 LRPDLKRGNTFEEDELLIKLHSLGKWSLIAGRLPGRTDNEIKYWNTHIKRKLISRG 120
Db 61 LRPDLKRGNTFEEDELLIKLHSLGKWSLIAGRLPGRTDNEIKYWNTHIKRKLISRG 120
OY 121 IDPQTHRLPNOTANTNTVTAPELDFRNSPTSVSKSSIKRPSLDENFNEOFKNTDSL 176
Db 121 IDPQTHRLPNOTANTNTVTAPELDFRNSPTSVSKSSIKRPSLDENFNEOFKNTDSL 176
OY 177 TSLLEPNCTASSGMTTDEEOQLHKKOQYGPNSGODINLELSIGIVSADSSRYSNANS 236

```

```

Db 171 SSSDDTKCNSS---TTEESQSLITPPPKKEEKSVPLVDLSLGLPQSQCNKSVSLN 226
OY 237 AESKPKVDNNNFQFLDAQAMAKAVCLCWLQGFSTSEICRNCNSNSNGFYSICR 289
Db 227 SSSSGFYD---LFRPPAKVAKRCVCCKWTLGJLQXGDEPCNCO--SFNGFYRIC 274

```

## RESULT 3

```

O9S9K9 PRELIMINARY; PRT; 257 AA.
ID O9S9K9;
DT 01-MAY-2000 (TREMBlrel. 13, Created)
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)
DE T22J18.19 PROTEIN.
GN T22J18.19.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eucosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RA Vysotskaya V.S., Schwartz J.R., Toriumi M., Yu G., Oji O., Kwan A.,
RA Liu S., Li J., Araujo R., Au M., Brendel V., Buehler E., Conway A.B.,
RA Conway A.R., Dewar K., Feng J., Kim C., Kurtz D., Li Y., Palm C.J.,
RA Shih P., Sun H., Davis R.W., Ecker J.R., Federspiel N.A.,
RA Theologis A.;
RA "Arabidopsis thaliana chromosome 1 BAC T22J18 sequence, complete
RT sequence.";
RT Submitted (JAN-1998) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA Theologis A.;
RA Submitted (JUN-1998) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RA Theologis;
RL Submitted (JUL-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL: AC003879; AAC25522.1; -.
DR HSSP: P01103; IPOK.
DR InterPro: IPR001005; Myb_DNA_bind.
DR Pfam: PF00249; myb_DNA-binding; 2.
DR SMART: SM00395; SANT; 2.
DR PROSITE: PS00037; MYB_1; UNKNOWN_1.
DR PROSITE: PS00334; MYB_2; 1.
DR PROSITE: PS00090; MYB_3; 2.
SQ SEQUENCE 257 AA; 29352 MW; 132D80DEA58D7D93 CRC64;

```

Query Match 48.5%; Score 768; DB 10; Length 257;  
 Best Local Similarity 55.6%; Pred. No. 2.8e-59;  
 Matches 160; Conservative 33; Mismatches 53; Indels 42; Gaps 7;

```

OY 1 MGRSPCCERKANTNKGAWKEDQRLINIRVHGECWRSLEPKAAGLLRCGSKRLRWNY 60
Db 1 MGRSPCCERKANTNKGAWKEDQRLINIRVHGECWRSLEPKAAGLLRCGSKRLRWNY 60
OY 61 LRPDLKRGNTFEEDELLIKLHSLGKWSLIAGRLPGRTDNEIKYWNTHIKRKLISRG 120
Db 61 LRPDLKRGNTFEEDELLIKLHSLGKWSLIAGRLPGRTDNEIKYWNTHIKRKLISRG 120
OY 121 IDPQTHRLPNOTANTNTVTAPELDFRNSPTSVSKSSIKRPSLDENFNEOFKNTDSL 180
Db 121 IDPQTHRLPNOTANTNTVTAPELDFRNSPTSVSKSSIKRPSLDENFNEOFKNTDSL 180
OY 181 EEPNCTASSGMTTDEEOQLHKKOQYGPNSGODINLELSIGIVSADSSRYSNANS 233
Db 181 EEPNCTASSGMTTDEEOQLHKKOQYGPNSGODINLELSIGIVSADSSRYSNANS 233
OY 234 AESKPKVDNNNFQFLDAQAMAKAVCLCWLQGFSTSEICRNCNSNSNGFYSICR 209

```

DB 210 VDSASTRMWSELF-----GAGESDAVCLCCLRICLFRNESCRCRVS 253

RESULT 4

ID Q38851 PRELIMINARY; PRT; 236 AA.

AC Q38851; 01-NOV-1996 (TREMELREL. 01, Created)

DT 01-NOV-1996 (TREMELREL. 01, Last sequence update)

DE 01-JUN-2001 (TREMELREL. 17, Last annotation update)

DN DNA-BINDING PROTEIN.

GN MYB OR MYB2.

OS Arabidopsis thaliana (Mouse-ear cress).

OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;

OC eucosids II; Brassicales; Brassicaceae; Arabidopsids.

OX NCBI\_Taxid:3702;

RA SEQUENCE FROM N.A.

RC STRAIN-LANDSBERG. ERECTA;

RA MEDLINE-96132152; PubMed-8580966;

RL I.S.F., Parish R.W.;

RT "Isolation of two novel myb-like genes from Arabidopsis and studies on the DNA-binding properties of their products."

RL Plant J. 8:963-972(1995).

RN [2]

RP SEQUENCE FROM N.A.

RC STRAIN-LANDSBERG. ERECTA;

RA Headlewood J.L.;

RL Submitted (MAY-1995) to the EMBL/GenBank/DBJ databases.

RN [3]

RP SEQUENCE FROM N.A.

RA Spiegel L.A., Huang E.N., Nascimento L.U., de la Bastide M., Vill D.M.,

RA Preston R.R., Matero A., Shah R., O'Shaughnessy A., Rodriguez M.,

RA Sheher M., Schuch K., See L.H., Swaby I., Habermann K., Dedhia N.N.,

RA Mewes H.W., Lemcke K., Mayer K.F.X.;

RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.

RN [4]

RP SEQUENCE FROM N.A.

RA EU Arabidopsis sequencing project;

RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.

DR EMBL; U26936; AAA98761.1;

DR EMBL; AL161515; CAB78069.1;

DR HSSP; P01103; IPOM.

DR TRANSFAC; T02539;

DR Mendel; 6886; Atatrh; Myb; 6886.

DR InterPro; IPR001395; Aldc\_ket\_red.

DR Pfam; PF00249; myb\_DNA\_bind.

DR SMART; SM00395; SANT; 2.

DR PROSITE; PS00063; ALDOXERO\_REDUCTASE\_3; UNKNOWN\_1.

DR PROSITE; PS00037; MYB\_1; UNKNOWN\_1.

DR PROSITE; PS00334; MYB\_2; 1.

DR PROSITE; PS50090; MYB\_3; 2.

DR DNA-binding.

SO SEQUENCE 236 AA; 26473 MW; D53F9AC798D7B8CB CRC64;

Query Match 48.1%; Score 762; DB 10; Length 236;

Best Local Similarity 58.9%; Pred. No. 8.4e-59;

Matches 145; Conservative 39; Mismatches 40; Indels 22; Gaps 5;

QY 1 MGRSPCCERANTNKAMTKEEDORLINTYRVHGGCWRSLPKAGLLRCGKSCRLRWNY 60

DB 1 MGRSPCCERANTNKAMTKEEDORLINTYRVHGGCWRSLPKAGLLRCGKSCRLRWNY 60

QY 61 LRPDLKRGNTFEEDDELLIKLSLGNKWSLTAGRLPGRTDNEIKYNTHTIKRLISRG 120

DB 61 LRPDLKRGNTFEEDDELLIKLSLGNKWSLTAGRLPGRTDNEIKYNTHTIKRLISRG 120

QY 121 IDPQTHRLNQTANTNTVTAPELDFRNSPTSVSKSSINPSLDFNYNEQFQKSN 179

DB 121 IDPQTHRLNQTANTNTVTAPELDFRNSPTSVSKSSINPSLDFNYNEQFQKSN 179

DB 121 IDPQTHRLNQTANTNTVTAPELDFRNSPTSVSKSSINPSLDFNYNEQFQKSN 179

QY 180 LEEPNCASSGATTEDEQOLHKROQYSPNSQ-DINLELSIGIVSADSSRVSNANSAE 238

DB 164 SSD-NGASTSCITTEDLQ--NGECYSNDSGHINKMLDTLFGSGMSGRVGVSSAD 220

QY 239 SKPKVD 244

DB 221 SKPMCD 226

QY 239 SKPKVD 244

DB 221 SKPMCD 226

QY 239 SKPKVD 244

DB 221 SKPMCD 226

QY 239 SKPKVD 244

DB 221 SKPMCD 226

QY 239 SKPKVD 244

DB 221 SKPMCD 226

QY 239 SKPKVD 244

DB 221 SKPMCD 226

QY 239 SKPKVD 244

DB 221 SKPMCD 226

QY 239 SKPKVD 244

DB 221 SKPMCD 226

QY 239 SKPKVD 244

DB 221 SKPMCD 226

QY 239 SKPKVD 244

DB 221 SKPMCD 226

QY 239 SKPKVD 244

DB 221 SKPMCD 226

QY 239 SKPKVD 244

DB 221 SKPMCD 226

QY 239 SKPKVD 244

DB 221 SKPMCD 226

QY 239 SKPKVD 244

DB 221 SKPMCD 226

QY 239 SKPKVD 244

DB 221 SKPMCD 226

QY 239 SKPKVD 244

DB 221 SKPMCD 226

QY 239 SKPKVD 244

DB 221 SKPMCD 226

QY 239 SKPKVD 244

DB 221 SKPMCD 226

QY 239 SKPKVD 244

DB 221 SKPMCD 226

QY 239 SKPKVD 244

DB 221 SKPMCD 226

QY 239 SKPKVD 244

DB 221 SKPMCD 226

QY 239 SKPKVD 244

DB 221 SKPMCD 226

QY 239 SKPKVD 244

DB 221 SKPMCD 226

QY 239 SKPKVD 244

DB 221 SKPMCD 226

QY 239 SKPKVD 244

DB 221 SKPMCD 226

Query Match 47.8%; Score 756.5; DB 10; Length 251;

Best Local Similarity 50.7%; Pred. No. 2.7e-58;

Matches 152; Conservative 32; Mismatches 55; Indels 61; Gaps 6;

QY 1 MGRSPCCERANTNKAMTKEEDORLINTYRVHGGCWRSLPKAGLLRCGKSCRLRWNY 60

DB 1 MGRSPCCERANTNKAMTKEEDORLINTYRVHGGCWRSLPKAGLLRCGKSCRLRWNY 60

QY 61 LRPDLKRGNTFEEDDELLIKLSLGNKWSLTAGRLPGRTDNEIKYNTHTIKRLISRG 120

DB 61 LRPDLKRGNTFEEDDELLIKLSLGNKWSLTAGRLPGRTDNEIKYNTHTIKRLISRG 120

QY 121 IDPQTHRLNQTANTNTVTAPELDFRNSPTSVSKSSINPSLDFNYNEQFQKSN 177

DB 121 IDPQTHRLNQTANTNTVTAPELDFRNSPTSVSKSSINPSLDFNYNEQFQKSN 177

DB 121 IDPQTHRLNQTANTNTVTAPELDFRNSPTSVSKSSINPSLDFNYNEQFQKSN 177

```

Db 121 MDPHTHRLTAVIDAAAPTRPAQIAVAPARAAPTMTALPTKQOOPPVPESSDDGSSGA 180
QY 178 DSELPNCTASSGMTTDEEOQOLHKQOYGPSNGODINLELSIGIYSANSSRVNANSA 237
Db 181 TSTGPRC-----PDLNLDLSVGPAPADP----- 205
QY 238 ESKPVDNNNFQFLQAMVAKAVCLCMQJLFGTSEICRNCONSNS-----NGFYSYCRPLD 293
Db 206 -----TSHPYCLCRHLGFRGEGAC-SCROADSASQSGGFRFRLE 246

RESULT 6
043525 PRELIMINARY; PRT; 273 AA.
AC 043525;
DT 01-NOV-1996 (TREMBlrel. 01, Created)
DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)
DE TRANSCRIPTION FACTOR.
GN MYB OR THM27.
OS Lycopersicon esculentum (Tomato).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC Asteridae; euasterids I; Solanales; Solanaceae; Solanum.
OX NCBI_TaxID=4081;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. ALISA CRAIG; TISSUE=HYPOCOTYL;
RX MEDLINE=96270378; PubMed=8639738;
RA Llan O., Hamilton W.D.O., Merryweather A.;
RT Cloning and initial characterization of 14 myb-related cDNAs from
RT tomato (Lycopersicon esculentum cv. Alisa Craig).";
RL Plant Mol. Biol. 30:1009-1020(1996).
DR EMBL; X95296; CAA64614.1;
DR HSSP; P01103; IPOM.
DR TRANSFAC; T02886;
DR Mendel; 8912; Lyces; myb.8912.
DR InterPro; IPR01005; myb_DNA_bind.
DR Pfam; PF00249; myb_DNA-binding; 2.
DR SMART; SM00395; SANT; 2.
DR PROSITE; PS00037; MYB_1; 1.
DR PROSITE; PS00334; MYB_2; 1.
DR PROSITE; PS00090; MYB_3; 2.
DR PROSITE; PS00090; MYB_3; 2.
SO SEQUENCE 273 AA; 31100 MW; 6F2E9365F3003AFB CRC64;

Query Match 46.0%; Score 728; DB 10; Length 273;
Best Local Similarity 53.0%; Pred. No. 9,2e-56;
Matches 151; Conservative 26; Mismatches 64; Indels 44; Gaps 5;

```

```

ID 09ATD1 PRELIMINARY; PRT; 264 AA.
AC 09ATD1;
DT 01-JUN-2001 (TREMBlrel. 17, Created)
DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)
DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)
DE GHMYB9.
GN GHMYB9.
OS Gossypium hirsutum (Upland cotton).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC euosids II; Malvales; Malvaceae; Gossypium.
OX NCBI_TaxID=3635;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ACALA MAXXA; TISSUE=DAY OF ANTHESIS OVULE;
RA Matz E.C., Burr B.;
RT "Cotton seed fibers are trichomes.";
RL Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF336286; AAK19619.1;
SO SEQUENCE 264 AA; 29629 MW; E09BCEBAE7662E95 CRC64;

Query Match 45.8%; Score 726; DB 10; Length 264;
Best Local Similarity 50.8%; Pred. No. 1.3e-55;
Matches 153; Conservative 27; Mismatches 53; Indels 68; Gaps 7;

```

```

QY 1 MGRSPCCERKANTKGAATKEEDORLINYIRVHGECGRSLPKAAGLRCCKSRLRMVY 60
Db 1 MGRSPCCERKANTKGAATKEEDORLINYIRVHGECGRSLPKAAGLRCCKSRLRMVY 60
QY 61 LRPLDKGNFTEDEDELITKLSLGNKWSLIAGRLPGRTDNEIKNYNTHIRKLLISRG 120
Db 61 LRPLDKGNFTEDEDELITKLSLGNKWSLIAGRLPGRTDNEIKNYNTHIRKLLISRG 120
QY 121 IDPQTHRLPNOTANTNTVTAPELDFRNSPTSVSKSSSIKNPDLRYNNEFOFKSTDSL 180
Db 121 IDPQTHRLPNOTANTNTVTAPELDFRNSPTSVSKSSSIKNPDLRYNNEFOFKSTDSL 180
QY 121 IDPQTHRLPNOTANTNTVTAPELDFRNSPTSVSKSSSIKNPDLRYNNEFOFKSTDSL 180
Db 121 IDPQTHRLPNOTANTNTVTAPELDFRNSPTSVSKSSSIKNPDLRYNNEFOFKSTDSL 180
QY 181 EEPNCTASSGMTTDEEOQOLHKQOYGPSNGODINLELSIGIYSANSSRVNANSAESK 240
Db 181 EEPNCTASSGMTTDEEOQOLHKQOYGPSNGODINLELSIGIYSANSSRVNANSAESK 240
QY 158 EMINIAAEFVETSKESDNNEI--IOEKSSSCLPLDLNLELRI-----SPHH 201
Db 158 EMINIAAEFVETSKESDNNEI--IOEKSSSCLPLDLNLELRI-----SPHH 201
QY 241 PKVDNNNFQFLQAMVAKAVCLCMQJLFGTSEICRNCONSNSNGF 285
Db 202 QOLDHHRHH-----QNSSSLCTFCISLGKSKDCSGSSSGNGM 241

```

```

RESULT 8
09S2P1 PRELIMINARY; PRT; 282 AA.
ID 09S2P1;
AC 09S2P1;
DT 01-MAY-2000 (TREMBlrel. 13, Created)
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)
DE POTATIVE TRANSCRIPTION FACTOR (MYB4).
GN F20M13.180 OR AT4G38620.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC euosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=K. F. X.; Kutzner M., Wambutt R., Bancroft I., Mewes H.W.,
RA Beyer K.F.X., Schueller C.;
RL Submitted (FEB-1999) to the EMBL/GenBank/DBJ databases.
SO SEQUENCE FROM N.A.
EU Arabidopsis sequencing project;

```

RL Submitted (FEB-1999) to the EMBL/GenBank/DBJ databases.  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RA Murphy G., Ridley P., Hudson S., Mewes H.W., Lemcke K., Mayer K.F.X.;  
 RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.  
 RN [4]  
 RP SEQUENCE FROM N.A.  
 RA EU Arabidopsis sequencing project;  
 RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: AL035540; CAB37518.1;  
 DR EMBL: AL161593; CAB80526.1;  
 DR HSSP: P01103; IPOM.  
 DR InterPro: IPR001005; MYB\_DNA\_bind.  
 DR Pfam: PF00249; MYB\_DNA-binding; 2.  
 DR SMART: SM00395; SANT; 2.  
 DR PROSITE: PS00037; MYB\_1; 1.  
 DR PROSITE: PS00334; MYB\_2; 1.  
 DR PROSITE: PS50090; MYB\_3; 2.  
 DR PROSITE: PS50090; MYB\_3; 2.  
 SQ SEQUENCE 282 AA; 31808 MW; CEFBAFIE9D8A8687 CRC64;

Query Match: 45.6%; Score 723; DB 10; Length 282;  
 Best Local Similarity 52.0%; Pred. No. 2.6e-55;  
 Matches 146; Conservative 38; Mismatches 49; Indels 48; Gaps 7;

QY 1 MGRSPCEKATNKGAATKEEDORLINTYRVHGGCGMRLPKAAGLLRCGKSCRLRWNY 60  
 DB 1 MGRSPCEKATNKGAATKEEDORLINTYRVHGGCGMRLPKAAGLLRCGKSCRLRWNY 60  
 QY 61 LRPDLKRGNTFEEDDELLIKLHSLGNKWSLIAGRLPGRTDNEIKYWNTHIRKRLISRG 120  
 DB 61 LRPDLKRGNTFEEDDELLIKLHSLGNKWSLIAGRLPGRTDNEIKYWNTHIRKRLISRG 120  
 QY 121 IDPOTHRPLNQTANTYVTAPELDFRNSPT-SYSSSSSTKNSLDFENYEFQKSTDS 179  
 DB 121 IDPOTHRPLNQTANTYVTAPELDFRNSPT-SYSSSSSTKNSLDFENYEFQKSTDS 179  
 QY 121 IDPOTHRPLNQTANTYVTAPELDFRNSPT-SYSSSSSTKNSLDFENYEFQKSTDS 176  
 DB 121 IDPOTHRPLNQTANTYVTAPELDFRNSPT-SYSSSSSTKNSLDFENYEFQKSTDS 176  
 QY 180 LEPNCTASGWTDEEOQLKKOQYGPNSGODINLEL----- 219  
 DB 177 I-----SMLETKKEKDECPVOEKF-----PDNLLELRISLPDDVDRLQHGKSTTP 222  
 QY 220 -----SIGIVSADSSRVSNAN-----SAESKPKVDNNNFQFL 251  
 DB 223 RCFKCSLGMINGMECRGMRCDYVGSSKSGSDMSNGDFL 263

RESULT 9  
 ID 049774 PRELIMINARY; PRT; 282 AA.  
 AC 049774;  
 DT 01-JUN-1998 (TREMBLrel. 06, Created)  
 DT 01-JUN-1998 (TREMBLrel. 11, Last sequence update)  
 DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)  
 DE PUTATIVE TRANSCRIPTION FACTOR (R2R3-MYB TRANSCRIPTION FACTOR).  
 GN MYB4.  
 OS Arabidopsis thaliana (Mouse-ear cress).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
 OC eurosids II; Brassicales; Brassicaceae; Arabidopsids.  
 OX NCBI\_TaxID=3702;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-CV. COLUMBIA;  
 RX MEDLINE=99056848; PubMed=9839469;  
 RA Kitzan H.D., Denekamp M., Greco R., Jin H.-L., Leyva A., Meissner R.,  
 RA Petroni K., Uzaliqui A., Bevan M., Martin C., Smeekens S.,  
 RA Tonelli C., Paz-Ares J., Weishaar B.;  
 RT "Towards functional characterisation of the members of the R2R3-MYB  
 gene family from Arabidopsis thaliana.";  
 RL Plant J. 16:263-276(1998).  
 RN [2]  
 RP SEQUENCE OF 55-99 FROM N.A.  
 RC TISSUE-WHOLE PLANTS;

RA Romero I., Fuentes A., Benito M.J., Malpica J., Leyva A., Paz-Ares J.;  
 RL Submitted (MAY-1997) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: AF062860; AAC3582.1;  
 DR EMBL: Z95763; CAB09195.1;  
 DR HSSP: P01103; IPOM.  
 DR Mendel: 27725; Arab\_MyB; 27725.  
 DR InterPro: IPR001005; MYB\_DNA\_bind.  
 DR Pfam: PF00249; MYB\_DNA-binding; 2.  
 DR SMART: SM00395; SANT; 2.  
 DR PROSITE: PS00037; MYB\_1; 1.  
 DR PROSITE: PS00334; MYB\_2; 1.  
 DR PROSITE: PS50090; MYB\_3; 4.  
 SQ SEQUENCE 282 AA; 31820 MW; F4CB4F0B47B575AD CRC64;

Query Match: 45.3%; Score 717; DB 10; Length 282;  
 Best Local Similarity 51.6%; Pred. No. 8.7e-55;  
 Matches 145; Conservative 38; Mismatches 50; Indels 48; Gaps 7;

QY 1 MGRSPCEKATNKGAATKEEDORLINTYRVHGGCGMRLPKAAGLLRCGKSCRLRWNY 60  
 DB 1 MGRSPCEKATNKGAATKEEDORLINTYRVHGGCGMRLPKAAGLLRCGKSCRLRWNY 60  
 QY 61 LRPDLKRGNTFEEDDELLIKLHSLGNKWSLIAGRLPGRTDNEIKYWNTHIRKRLISRG 120  
 DB 61 LRPDLKRGNTFEEDDELLIKLHSLGNKWSLIAGRLPGRTDNEIKYWNTHIRKRLISRG 120  
 QY 121 IDPOTHRPLNQTANTYVTAPELDFRNSPT-SYSSSSSTKNSLDFENYEFQKSTDS 179  
 DB 121 IDPOTHRPLNQTANTYVTAPELDFRNSPT-SYSSSSSTKNSLDFENYEFQKSTDS 179  
 QY 121 IDPOTHRPLNQTANTYVTAPELDFRNSPT-SYSSSSSTKNSLDFENYEFQKSTDS 176  
 DB 121 IDPOTHRPLNQTANTYVTAPELDFRNSPT-SYSSSSSTKNSLDFENYEFQKSTDS 176  
 QY 180 LEPNCTASGWTDEEOQLKKOQYGPNSGODINLEL----- 219  
 DB 177 I-----SMLETKKEKDECPVOEKF-----PDNLLELRISLPDDVDRLQHGKSTTP 222  
 QY 220 -----SIGIVSADSSRVSNAN-----SAESKPKVDNNNFQFL 251  
 DB 223 RCFKCSLGMINGMECRGMRCDYVGSSKSGSDMSNGDFL 263

RESULT 10  
 ID P81393 PRELIMINARY; PRT; 232 AA.  
 AC P81393;  
 DT 01-JUN-1998 (TREMBLrel. 06, Created)  
 DT 01-JUN-1998 (TREMBLrel. 06, Last sequence update)  
 DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)  
 DE MYB-RELATED PROTEIN 308.  
 GN MYB 308.  
 OS Antirrhinum majus (Garden snapdragon).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
 OC Asteridae; euasterids I; Lamiales; Veroniceae; Antirrhinum.  
 OX NCBI\_TaxID=4151;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-JT:522; TISSUE=FLOWER BUDS;  
 RX MEDLINE=93005689; PubMed=1840903;  
 RA Jackson D., Cullanez-Macia F., Prescott A.G., Roberts K., Martin C.;  
 RT "Expression patterns of myb genes from Antirrhinum flowers.";  
 RL Plant Cell 3:115-125(1991).  
 CC -1- FUNCTION: MAY BE A TRANSCRIPTIONAL ACTIVATOR.  
 CC -1- TISSUE SPECIFICITY: ROOT, STEM, LEAF, SEED POD, IMMATURE FLOWER  
 AND MATURE FLOWER.  
 CC -1- SIMILARITY: BELONGS TO THE MYB FAMILY OF PROTEINS.  
 CC PIR: J00960; J00960.  
 DR HSSP: P01103; IPOM.  
 DR TRANSFAC: T02926;  
 DR Mendel: 24821; Antma\_MyB; 24821.  
 DR InterPro: IPR001005; MYB\_DNA\_bind.  
 DR Pfam: PF00249; MYB\_DNA-binding; 2.  
 DR SMART: SM00395; SANT; 2.  
 DR PROSITE: PS00037; MYB\_1; 1.

DR PROSITE; PS00334; MYB\_2; 1.  
 DR PROSITE; PS50090; MYB\_3; 2.  
 KW Nuclear protein; DNA-binding; Repeat; Transcription regulation.  
 FT DNA\_BIND 62 102 MYB.  
 FT DNA\_BIND 62 102 MYB.  
 SQ SEQUENCE 232 AA; 26164 MW; 8B3943FCB0D403D4 CRC64;

Query Match 44.5%; Score 704.5; DB 10; Length 232;  
 Best Local Similarity 52.4%; Pred. No. 8.5e-54;  
 Matches 150; Conservative 24; Mismatches 55; Indels 57; Gaps 8;

QY 1 MGRSPCCAKHTNKGAMTKEEDRLINIRVHGGCGRSLPKAAGLRCKGSCRLRWNY 60  
 DB 1 MGRSPCCAKHTNKGAMTKEEDRLVAVIRAHGGCGRSLPKAAGLRCKGSCRLRWNY 60  
 QY 61 LRPDLKRGNTFEEDDLIKLHSLGNKWSLAGRLPGRTDNEIKYWNTHIKRKLISRG 120  
 DB 61 LRPDLKRGNTFEEDDLIKLHSLGNKWSLAGRLPGRTDNEIKYWNTHIKRKLISRG 120  
 QY 121 IDPOTHRPLNO-TANTNTYTAPELDFRNSPTSVSKSSSTKNPSLDNFYNEQFSNNTDS 179  
 DB 121 IDPOTHRPLNO-TANTNTYTAPELDFRNSPTSVSKSSSTKNPSLDNFYNEQFSNNTDS 179  
 QY 180 LEEPCNCTASSGWTDEEOEOLHKKQOYGPSNGO--DINELSIGIVSADSSRYSNANSA 237  
 DB 150 KEE-----DTKHKAVNDIIRKENSFYQERCPDLNLDKI---SPCOQOQINHYOE 197  
 QY 238 ESKPKYDNNNFQLEQAMVAKAVCLQWLGFGTSEICRNQNSN 283  
 DB 198 NLKTGGRNG-----SSTLCFVCRLGIONSKDC-SCSDGVGN 232

RESULT 11  
 ID 0957Y2 PRELIMINARY; PRT; 246 AA.  
 AC 0957Y2  
 DT 01-MAY-2000 (TREMBLrel. 13, Created)  
 DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)  
 DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)  
 DE PUTATIVE DNA BINDING PROTEIN.  
 GN FLZK8.1 OR MYB3.  
 OS Arabidopsis thaliana (Mouse-ear cress).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
 OC euroids II; Brassicales; Brassicaceae; Arabidopsis.  
 OX NCBI\_TaxID=3702;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Federle N.A., Palm C.J., Conway A.B., Conn L., Hansen N.F.,  
 RA Altairi H., Araujo R., Huizar L., Kowley D., Buehler E., Dunn P.,  
 RA Gonzalez A., Krementseva I., Kim C., Lenz C., Li J., Liu S.,  
 RA Luros S., Schwartz J., Shinn P., Toriumi M., Vysotskaia V.S.,  
 RA Walker M., Yu G., Ecker J., Theologis A., Davis R.W.,  
 RL Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.  
 RT [2]  
 RT SEQUENCE FROM N.A.  
 RC STRAIN=CV. COLUMBIA;  
 RC MEDLINE=99056848; PubMed=9839469;  
 RA Kriatz H.D., Denekamp M., Greco R., Jin H.-L., Leyva A., Meissner R.,  
 RA Petroni K., Urzabou A., Bevan M., Martin C., Smeekens S.,  
 RA Tonelli C., Paz-Ares J., Weissshaar B.;  
 RT "Towards functional characterisation of the members of the R2R3-MYB  
 gene family from Arabidopsis thaliana.";  
 RL Plant J. 16:263-276(1998).  
 DR EMBL; AC006551; AAF1815.1; -;  
 DR EMBL; AF062859; AAC83581.1; -;  
 DR HSSP; P01103; 1POM.  
 DR InterPro; IPR001005; MYB\_DNA\_bind.  
 DR Pfam; PF00249; MYB\_DNA-binding; 2.  
 DR SMART; SM00395; SANT; 2.  
 DR PROSITE; PS00037; MYB\_1; UNKNOWN\_1.  
 DR PROSITE; PS00334; MYB\_2; 1.  
 DR PROSITE; PS00334; MYB\_2; 1.

DR PROSITE; PS50090; MYB\_3; 2.  
 SQ SEQUENCE 246 AA; 28151 MW; 2CA8A529153FF486 CRC64;

Query Match 44.3%; Score 702; DB 10; Length 246;  
 Best Local Similarity 54.0%; Pred. No. 1.5e-53;  
 Matches 149; Conservative 33; Mismatches 52; Indels 42; Gaps 7;

QY 13 NKGATKEEDQRLINIRVHGGCGRSLPKAAGLRCKGSCRLRWNYLRPDLKRGNTFE 72  
 DB 2 NKGATKEEDQLVDYIRKHGGCGRSLPRAAGLRCKGSCRLRWNYLRPDLKRGNTFE 61  
 QY 73 EEDDLIKLHSLGNKWSLAGRLPGRTDNEIKYWNTHIKRKLISRGIDPOTHRPLNQT 132  
 DB 62 EEDDLIKLHSLGNKWSLAGRLPGRTDNEIKYWNTHIKRKLISRGIDPNSHRLINE 120  
 QY 133 ANNTNTYTAPELDFRNSPTSVSKSSSTKNPSLDNFYNEQFSNNTDSLEEP-----NCTA 187  
 DB 121 -----SVSPSSLDNDVYETLHLDSPGVKPPVPEEIGMANN-C-E 160  
 QY 188 SSGMTDEEOEOLHKKQOYGPSNGQDINELSIG-IVSADSSR-VSNANSAESKPKVDN 245  
 DB 161 SSGTTSKED-----YNEEDWYLNLELSVGPSTYESTRKAYSDASSTRWMS 210  
 QY 246 NNFQLEQAMVAKAVCLQWLGFGTSEICRNQNSN 281  
 DB 211 ELF-----GAHESDAVCLCGRIGLFRNESCRCNCRVSD 242

RESULT 12  
 ID 09SM28 PRELIMINARY; PRT; 274 AA.  
 AC 09SM28  
 DT 01-MAY-2000 (TREMBLrel. 13, Created)  
 DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)  
 DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)  
 DE DNA-BINDING PROTEIN.  
 GN MYB32.  
 OS Arabidopsis thaliana (Mouse-ear cress).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
 OC euroids II; Brassicales; Brassicaceae; Arabidopsis.  
 OX NCBI\_TaxID=3702;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA STRAIN=LANDSBERG ERECTA;  
 RA Li S.F., Heazlewood J., Parish R.W.,  
 RT "Cloning of three MYB-like genes from Arabidopsis thaliana (Accession  
 RT Nos. U26933, AF048841, and U26934) (FGR99-138).";  
 RL Plant Physiol. 121:313-313(1999).  
 DR EMBL; U26933; AAF13100.1; -;  
 DR HSSP; P01103; 1POM.  
 DR InterPro; IPR001395; Aldo\_ket\_red.  
 DR InterPro; IPR001005; MYB\_DNA\_bind.  
 DR Pfam; PF00249; MYB\_DNA-binding; 2.  
 DR SMART; SM00395; SANT; 2.  
 DR PROSITE; PS00063; ALDO-KETO\_REDUCTASE\_3; UNKNOWN\_1.  
 DR PROSITE; PS00037; MYB\_1; 1.  
 DR PROSITE; PS00334; MYB\_2; 1.  
 DR PROSITE; PS50090; MYB\_3; 2.  
 KW DNA-binding.  
 SQ SEQUENCE 274 AA; 31465 MW; 6C06ACB15D794CAC CRC64;

Query Match 43.2%; Score 684; DB 10; Length 274;  
 Best Local Similarity 49.5%; Pred. No. 6.3e-52;  
 Matches 146; Conservative 30; Mismatches 61; Indels 58; Gaps 9;

QY 1 MGRSPCCAKHTNKGAMTKEEDRLINIRVHGGCGRSLPKAAGLRCKGSCRLRWNY 60  
 DB 1 MGRSPCCAKHTNKGAMTKEEDKLIYIRAHGGCGRSLPRAAGLRCKGSCRLRWNY 60  
 QY 61 LRPDLKRGNTFEEDDLIKLHSLGNKWSLAGRLPGRTDNEIKYWNTHIKRKLISRG 120

Db 61 LRPDLKRGNTLEEDDLITKLHSLGNKWSLIATRLPGRTDNEIKYWNTHVKRLLKRG 120  
 QY 121 IDPOTHRPLNONTANTNTVTAPELDFRNSPTSVSKSSSIKPSLDENYNEOFKSNDSL 180  
 Db 121 IDPOTHRPLNONTANTNTVTAPELDFRNSPTSVSKSSSIKPSLDENYNEOFKSNDSL 158  
 QY 181 EEPNCTASSGNTDEEQOELHKKOQYGPSNQ--DINLELSIGIVSADSSRVSNANSAE 238  
 Db 159 EK--IANFG-----DERIQKREYSVVEERCLDNLLELRISPPMODKLFH----- 200  
 QY 239 SKPKVDNNNFOLFQAMAVAKVCLCMQLGFTSEICRNCN-----SNSNGFYS 287  
 Db 201 -----DERNLRFGR-----VKHRCACRFGFGNGKEC-SCNNVKQOTEDSSSSSYS 245

RESULT 13

049608 PRELIMINARY: PRT: 274 AA.  
 ID 049608  
 AC 049608  
 DT 01-JUN-1998 (TREMBlrel. 06, Created)  
 DT 01-JUN-1998 (TREMBlrel. 06, Last sequence update)  
 DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)  
 DE MYB-LIKE PROTEIN (MYB-LIKE PROTEIN).  
 GN MYB OR MAE13.50 OR AY4634950.  
 OS Arabidopsis thaliana (Mouse-ear cress).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
 OC eucosids II; Brassicales; Brassicaceae; Arabidopsids.  
 OX NCBI\_TaxID=3702;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Bevan M.F., Purnelle B., Boutry M., Goffeau A., Hohensel J., Mewes H.W.,  
 RA Mayer K.F.X., Schueller C.;  
 RL Submitted (MAR-1998) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RA Murphy G., Ridley P., Hudson S., Mewes H.W., Lemcke K., Mayer K.F.X.;  
 RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RA Purnelle B., Boutry M., Goffeau A., Mewes H.W., Lemcke K.,  
 RA Mayer K.F.X.;  
 RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.  
 RN [4]  
 RP SEQUENCE FROM N.A.  
 RA EU Arabidopsis sequencing project;  
 RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AL022023; CA117764.1; -  
 DR EMBL; AL161586; CAB80216.1; -  
 DR HSP; P01103; IPOM.  
 DR Mendel; 28225; Arabid.MYB.28225.  
 DR InterPro; IPR001395; AIGO\_kel\_red.  
 DR InterPro; IPR001005; MYB\_DNA\_bind.  
 DR Pfam; PF00249; MYB\_DNA\_bind.  
 DR SMART; SM00395; SANT; 2.  
 DR PROSITE; PS00063; ALDOKEO\_REDCTASE\_3; UNKNOWN\_1.  
 DR PROSITE; PS00037; MYB\_1; 1.  
 DR PROSITE; PS00334; MYB\_2; 1.  
 DR PROSITE; PS50090; MYB\_3; 2.  
 SQ SEQUENCE 274 AA; 31457 MW; 817C0609F7DB5517 CRC64;

Query Match 43.28; Score 684; DB 10; Length 274;  
 Best Local Similarity 49.58; Pred. No. 6.3e-52;  
 Matches 146; Conservative 30; Mismatches 61; Indels 58; Gaps 9;

QY 1 MGRSPCCERKATHTKGAATKEEDQRLINIRYHGECSWRLPRAAGLLRCGKSCRLLRWNY 60  
 Db 1 MGRSPCCERKATHTKGAATKEEDQRLINIRYHGECSWRLPRAAGLLRCGKSCRLLRWNY 60  
 QY 61 LRPDLKRGNTLEEDDLITKLHSLGNKWSLIAGRLPGRTDNEIKYWNTHVKRLLKRG 120

Db 61 LRPDLKRGNTLEEDDLITKLHSLGNKWSLIATRLPGRTDNEIKYWNTHVKRLLKRG 120  
 QY 121 IDPOTHRPLNONTANTNTVTAPELDFRNSPTSVSKSSSIKPSLDENYNEOFKSNDSL 180  
 Db 121 IDPOTHRPLNONTANTNTVTAPELDFRNSPTSVSKSSSIKPSLDENYNEOFKSNDSL 158  
 QY 181 EEPNCTASSGNTDEEQOELHKKOQYGPSNQ--DINLELSIGIVSADSSRVSNANSAE 238  
 Db 159 EK--IANFG-----DERIQKREYSVVEERCLDNLLELRISPPMODKLFH----- 200  
 QY 239 SKPKVDNNNFOLFQAMAVAKVCLCMQLGFTSEICRNCN-----SNSNGFYS 287  
 Db 201 -----DERNLRFGR-----VKHRCACRFGFGNGKEC-SCNNVKQOTEDSSSSSYS 245

RESULT 14

023889 PRELIMINARY: PRT: 239 AA.  
 ID 023889  
 AC 023889  
 DT 01-JAN-1998 (TREMBlrel. 05, Created)  
 DT 01-JAN-1998 (TREMBlrel. 05, Last sequence update)  
 DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)  
 DE OSMYB1.  
 GN OSMYB1.  
 OS Oryza sativa (Rice).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
 OC Ehrhartoideae; Oryzaceae; Oryza.  
 OX NCBI\_TaxID=4530;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=SEED;  
 RX MEDLINE=98036074; PubMed=9370307;  
 RA Suzuki A., Suzuki T., Tanabe F., Toki S., Washida H., Wu C.-Y.,  
 RA Takaiwa F.;  
 RL "Cloning and expression of five myb-related genes from rice seed."  
 RL Gene 198:393-398(1997).  
 DR EMBL; D88617; BAA23337.1; -  
 DR HSP; P01103; IPOM.  
 DR TRANSFAC; T02869; -  
 DR Mendel; 24458; Oryza.MYB.24458.  
 DR InterPro; IPR001005; MYB\_DNA\_bind.  
 DR Pfam; PF00249; MYB\_DNA\_binding; 2.  
 DR SMART; SM00395; SANT; 2.  
 DR PROSITE; PS00037; MYB\_1; UNKNOWN\_1.  
 DR PROSITE; PS00334; MYB\_2; 1.  
 DR PROSITE; PS50090; MYB\_3; 2.  
 SQ SEQUENCE 239 AA; 26725 MW; B937E5E7848826F4 CRC64;

Query Match 43.18; Score 683; DB 10; Length 239;  
 Best Local Similarity 93.08; Pred. No. 6.6e-52;  
 Matches 120; Conservative 7; Mismatches 2; Indels 0; Gaps 0;

QY 1 MGRSPCCERKATHTKGAATKEEDQRLINIRYHGECSWRLPRAAGLLRCGKSCRLLRWNY 60  
 Db 1 MGRSPCCERKATHTKGAATKEEDQRLINIRYHGECSWRLPRAAGLLRCGKSCRLLRWNY 60  
 QY 61 LRPDLKRGNTLEEDDLITKLHSLGNKWSLIAGRLPGRTDNEIKYWNTHVKRLLKRG 120  
 Db 61 LRPDLKRGNTLEEDDLITKLHSLGNKWSLIAGRLPGRTDNEIKYWNTHVKRLLKRG 120  
 QY 121 IDPOTHRPL 129  
 Db 121 IDPOTHRPL 129

RESULT 15  
 ID 042379 PRELIMINARY: PRT: 269 AA.  
 AC 042379  
 DT 01-NOV-1996 (TREMBlrel. 01, Created)  
 DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)

Search completed: January 23, 2002, 21:21:03  
 Job time: 565 sec

01-JUN-2001 (TRENBLREL, 17, last annotation update)  
 DE Y49 GENE.  
 GN MYB OR Y49 OR T24121.13.  
 OS Arabidopsis thaliana (Mouse-ear cress).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
 OC eurosids II; Brassicales; Brassicaceae; Arabidopsids.  
 NCBI\_TaxID=3702;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Quedvlieg N.E.M., Dockx J., Keultjes G.G.M., Smeekens J.C.M.;  
 RL Submitted (Aug-1995) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-LANDSBERG ERECTA;  
 RX MEDLINE=96132152; PubMed=8580966;  
 RA Li S.F., Parish R.W.;  
 RT Isolation of two novel myb-like genes from Arabidopsis and studies on  
 the DNA-binding properties of their products.\*;  
 RL Plant J. 8:963-972(1995).  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-LANDSBERG ERECTA;  
 RA Heazlewood J.L.;  
 RL Submitted (May-1995) to the EMBL/GenBank/DBJ databases.  
 RN [4]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-CV. COLUMBIA;  
 RA Lin X., Kaul S., Shea T.P., Fujii C.Y., Shen M., Vanaken S.E.,  
 Barnstead M.E., Mason T.M., Bowman C.L., Rongning C.M., Benito M.,  
 Carrera A.J., Creasy T.H., Buell C.R., Town C.D., Nierman W.C.;  
 RA Fraser C.M., Venter J.C.;  
 RT Arabidopsis thaliana chromosome II BAC T24121 genomic sequence.\*;  
 RL Submitted (Apr-1999) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; X90385; CAA62033.1; -;  
 DR EMBL; U26937; AAA98762.1; -;  
 DR EMBL; AC005825; AAD24605.1; -;  
 DR HSSP; P01103; 1POM.  
 DR TRANSFAC; T02540; -;  
 DR Mendel; 6896; Arabidopsis; MYB\_DNA\_bind.  
 DR InterPro; IPR001005; MYB\_DNA\_bind.  
 DR Pfam; PF00249; MYB\_DNA\_binding; 2.  
 DR SMART; SM00395; SANT; 2.  
 DR PROSITE; PS00037; MYB\_1; 1.  
 DR PROSITE; PS00334; MYB\_2; 1.  
 DR PROSITE; PS50090; MYB\_3; 2.  
 KW DNA-binding.  
 SO SEQUENCE 269 AA; 31009 MW; 7A6C5D7C3CA0135C CRC64;

Query Match 43.0%; Score 681.5; DB 10; Length 269;  
 Best Local Similarity 48.8%; Pred. No 1e-51;  
 Matches 142; Conservative 35; Mismatches 59; Indels 55; Gaps 8;

OY 1 MGRSPCCERAHNTNKGAMTKCEDORLNIYIRVHGCGKRSILPKAGLLRCGKSCRLRWINY 60  
 DB 1 MGRSPCCERAHNTNKGAMTKCEDORLNIYIRVHGCGKRSILPKAGLLRCGKSCRLRWINY 60  
 OY 61 LRPDLRKGNFTDEDELLIKLHSLGKMSLIAGRLPGRTDNEIKNYWNTHIKRLKLSKG 120  
 DB 61 LRPDLRKGNFTDEDELLIKLHSLGKMSLIAGRLPGRTDNEIKNYWNTHIKRLKLSKG 120  
 OY 121 IDPOTHRPLNOTANTVTATPELDFNPSPTSVKSSSIKNPSLDFNNEFOFKSNTDSL 180  
 DB 121 IDPOTHRPLNOTANTVTATPELDFNPSPTSVKSSSIKNPSLDFNNEFOFKSNTDSL 180  
 OY 181 EEPNCTASGAMTTDEQ---QEQLLHKQOYGPSNGODINLELSIGIVASADSSRVSNANSA 237  
 DB 163 OKONKTIKRNGLVCKEERVVVEEKI-----GPDNLNLELRISPPWQNRISTCTAS 212  
 OY 238 ESKPKYDNNNFOLFQAMAKAVCLCQWLGFTSEICRNCQNSNNGFYSY 288  
 DB 213 RF-----YMEQDMEC-----SSETYK-CQTESSSI-SY 239